Zhongxue kaoshixue gailun

# 中学考试学概论

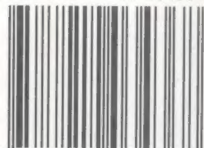
闫蒙钢 ● 著

安徽人民出版社

责任编辑：戴明芳
封面设计：王　芳

Zhongxue kaoshixue gailun

# 中学考试学概论



ISBN 978-7-212-03436-8

定价：26.00元

# 中学考试学概论

闫蒙钢 著

安徽人民出版社

**图书在版编目（CIP）数据**

中学考试学概论/闫蒙钢著．—合肥：安徽人民出版社，2008.12
ISBN 978－7－212－03436－8

Ⅰ．中…　Ⅱ．闫…　Ⅲ．中学—考试学—概论　Ⅳ．G632.474

中国版本图书馆 CIP 数据核字（2008）第 190324 号

**中学考试学概论**

**闫蒙钢　著**

出版发行：安徽人民出版社
地　　址：合肥市政务文化新区圣泉路 1118 号出版传媒广场 8 楼
发 行 部：0551－3533258　3533268　3533292（传真）　邮编：230071
组　　编：安徽师范大学编辑部　电话：0553－3883578　3883579
经　　销：新华书店
印　　制：芜湖新欣传媒有限公司
开　　本：889×1194　1/32　印张：11　字数：276 千
版　　次：2009 年 1 月第 1 版　2009 年 1 月第 1 次印刷
标准书号：ISBN 978－7－212－03436－8
定　　价：26.00 元

# 前　言

考试是人类社会特有的一种测量活动，是测度、甄别人的知识与智能个别差异的一种社会活动。它不仅是学校教学工作的重要环节，而且是现代人事管理的有力工具，具有服务于现代经济的重要社会职能。考试与教育是一对孪生子，产生于社会生产和社会生活的客观需要，并随社会发展需求的变化而不断改变其内容、性质和模式，形成与其社会相适应的结构，以发挥不同社会或时代所追求的功能。考试对社会的影响是多方面的，在社会和教育大变革时代，能否打造优质的考试，对社会的稳定与和谐都至关重要。当前，重要的是在传统考试基础上构建一个科学的教育考试与评价体系，实现从单一考试到多元评价的跃升，这既是考试改革的核心问题，也是本书编写所追求的目标。

中国素有“考试故乡”之称，中国的考试文明源远流长。在世界范围内教育改革飞速发展的新时期，研究在中学考试中如何正确发挥考试的督导功能，使之为亿万后继人才能力素质的全面发展导以正确方向；如何有效控制考试误差，使人才评价、甄选的依据准确可靠；如何形成反映考试活动运行规律的考试理论体系，不断丰富发展和普及考试设计、实施、评价、管理的科学知识，加快考试的科学化、规范化、现代化进程，也是我们教育工作者肩负的历史责任。

《中学考试学概论》以中学考试为研究对象，探求考试发生、发展、演变的成因，论述考试的基本理论、程序、原则与方法，

研究考试的内在结构、外部关系及其功能，揭示考试的固有规律，用以指导考试实践，并为考试科学的各分支学科奠定理论基础。

本书在确定编排体系与编选内容时坚持从中学教育实际出发，综合运用教育统计学、教育测量学、教育评价学和计算机科学的基本原理与技术，结合中学各学科特点和当前考试改革的热点，对我国基础教育课程改革中的学业成绩考试改革的特征和趋势进行了阐述；对科学化考试的基本理论、方法与技术作了系统的论述；对与中学教育密切相关、且影响重大的“中考、会考、高考和竞赛考试”作了专题评析；对中学生考试心理也进行了深入探讨。本书旨在引导读者掌握科学化考试的理论与方法，初步学会试题的编制技术，有效地组织管理考试，不断提高考生考试的心理素质和参加各类考试的能力，并期望能对我国考试学学科体系的构建和完善产生积极作用。

本书在编写时力求史论结合、中外兼顾的研究思路，注重以基础教育课程改革为平台，积极借鉴国内外考试史上的经验教训，在诸多方面有所创新，内容厚实，表现出新颖性、时代性、理论性、实践性、操作性和可读性等特点，便于教学与自学，具有一定的学术价值和实用价值。不仅适用于高师院校师范类专业教学使用，也可作为中学教师继续教育、教育科研人员、教育硕士、课程与教学论研究生学习研究的指导用书，还可用作报考教育类研究生的参考用书。

全书由闫蒙钢教授拟定基本框架和撰写提纲。具体分工为：第一章和第九章的第一、二节阙荣辉；第二章张艳华；第三、十二章张四方；第四、五、六、八、十章及第九章的第三节闫蒙钢；第十一章马旭明、许苗苗；第七章朱小丽；第十三章胡行忠；第十四章杨亚丽；第十五章张犟。张艳华、张犟做了书稿的整理工作，张画花、蒋业健、汤伶俐、陈斐参与了部分资料的搜集工作。全书由闫蒙钢教授统稿、审稿和定稿。在编写过程中，参阅和引

用了很多相关研究成果，在此，谨向有关专家和学者一并表示诚挚的谢意；同时也感谢安徽人民出版社安徽师范大学编辑部及本书的编辑所付出的辛勤劳动，感谢安徽师范大学给予的2008年度出版基金资助。

限于作者水平，书中一定还存在着不完善之处，敬请读者不吝指正。

闫蒙钢

2009年1月

# 目　录

# 第一章 中学考试学的任务

考试是中国人的一大发明。当人们需要测度、甄别人的知识和才能的差异以“选贤与能”的时候，考试便产生了。自此以后，考试便随着人类社会的兴衰更迭经历了人类文明发展的全程。如今“考试已是这样稳固地站定了脚跟，要废除它，似乎比取消篝火节或圣诞节更无可能”①，它已渗透于现实社会的各行各业，成为人们有效开发和科学利用社会人力资源、提高活动效率、促进社会发展的基本措施之一，并随着人与社会的发展而不断发展创新。在教育领域，考试更是具有不可替代的巨大影响力。

## 第一节 考试的意义

考试既是一个普遍的范畴，又是一个永恒的范畴②。凡是有从事物质生产活动和追求活动效率的地方，就必有测度、甄别、判断人的身心素质差异、评价人的活动效率高低的考试存在。考试是随着人类自身发展及社会发展的内在需求而产生，又在动态性满足人及社会需求的变化发展中而不断延伸，同时又不断地影响并促进着教育的发展。

---

① （英）罗伯特·蒙哥马利（Robet Montgomery）. 黄鸣译. 考试的新探索［M］. 南宁：广西人民出版社，1984：76.

② 廖平胜. 考试学原理［M］. 武汉：华中师范大学出版社，2003：2.

## 一、考试与人的发展

考试之所以问世于人类，是由人与外部世界的本质关系所决定的，即以考试敦促新生一代继承并发展已有的生活经验、劳动技能、意志品质等，使之能通过活动来取得一定的外界物，从而满足自己的需要。这表明考试活动在人们获得利益和满足需要等方面具有手段属性，是制约人的活动效率、影响人的活动目的实现的重要因素之一。

同时，考试是人的活动的知识性、依存性和目的性的客观需求，是实现人的社会化和人的社会价值的重要手段，也是人类社会演进发展不可或缺的机制。在社会人才系统中它既充当着育才与用才的中介，又与人才培养、人才使用的过程密切关联，是实现人的社会化、人力资源科学管理和人的价值需求的重要机制。考试产生与发展的历程，体现了人类进化的需求及发展的进程，其活动目的始终指向人的发展。

再者，考试活动的本质职能是测度、甄别、评价人自身素质的个别差异，是人类特有的自身评价活动，“人本化”是其显著特征。由于考试活动中的主、客体都是现实的、活动着的人，所以无论是考试主体设计、实施考试，还是考试客体参加应考，都是人出于对某种特定目的的追求，以满足某种个人的、群体的或社会的需要。

值得注意的是，考试活动之所以自其诞生以来一直伴随人类社会始终，对人的发展具有特定的现实意义，归根结底是因为考试活动是人的活动，是为了人的发展的活动。也就是说，服务人的发展，促进人的发展，实现人的发展，应该而且必须是考试活动的灵魂和考试活动的目的。

## 二、考试与社会发展

人类要想自己的优秀文化绵延不断、使新生一代获得与其所

处社会发展需求相适应的文化知识与技能，并负起推动社会向前的历史使命，就必须依靠教育。而要想圆满实现教育目标，有效组织、协调单个社会成员之间的关系及其他社会关系，考试这一育才与用才的社会中介不可或缺。尤其是近代社会的快速发展，经济竞争日趋激烈，人力资源的合理配置与有效使用是关键，而其基础又在于教育的健康持续发展和考试的科学化。因此，振兴教育和革新考试，便成为世界各国应对新技术革命挑战的首选措施。

### 三、考试与教育发展

考试自从其诞生以来，就一直深刻影响着教育及与教育相关的各类教育主体和教育客体，当然受其影响最直接的是学生。隋唐以来，许多考生因不能金榜题名而郁郁终生，也有人为“考而优则仕”欣喜若狂。时过境迁，现在的六月高考之后，经常是“几家欢喜几家愁”。但是，不管你给考试罗列多少罪状，有一点是不能否定的：它打破了门第和血缘的界限，而代之以分数录取人才，这无疑是文明的表现和历史的进步。在整个教育资源尚未达到“按需分配”的前提下，社会整体的发展要求最有效的利用这些资源，一种行之有效的方法是，采用一定的手段或方法来确定“精英”以享受数量有限的教育资源和优质的教育资源，而这种分辨的最好方法便是考试。因此，考试能有效促进教育的发展，反过来，教育的发展也能带动考试科学的进步。

由此可见，人、社会及教育发展对考试的需求是必然的。因此，自古以来考试——尤其是各行各业的合格与选拔性考试，就一直是人们争相角逐的目标和希望视线的焦点。

## 第二节　考试的本质和功能

考试活动源远流长，它不仅广泛运用于社会生活的各个方面，

而且也成为人们潜心探究的一个学科领域。然而，无论是考试科学理论体系的建构，还是考试实践模式的革旧创新，都与人们对考试本质与功能的认识程度密切相关。对考试本质属性概念的厘清，在很大程度上决定了科学的考试理论体系的建立；同样，对考试功能的清楚认识，就能较准确地把握考试实践发展的价值取向。

## 一、考试的本质

考试本质隐藏于繁复的考试现象之中，惟有透过纷繁的考试现象才能揭示考试的本质，所以，准确界定考试概念的内涵、科学分析考试活动的属性和考试关系，是把握考试本质的前提。

考试活动自其产生以来，在漫长的发展演变过程中，形成了众多反映自身特质的概念，而“考试”则是其中最基本的概念。由于其使用时间的久远和使用面的宽广，古今中外有关“考试”内涵的界定可谓仁智互见。如《教育大辞典》中的工具说[1]、《中国大百科全书》（教育）中的方法说[2]、《中国劳动人事百科全书》中的手段说[3]和活动说[4]等。这些观点各自从不同角度描述了考试的形成，并揭示了考试的某些特点和一般特性。

综合前人的研究和上述分析，我们认为，考试是一定组织中的考试主体根据考试目的的需要，选择运用有关资源，对考试客体某方面或诸方面的素质水平进行测度、甄别和评价的一种社会活动。

---

① 顾明远．教育大辞典（第一卷）［M］．上海：上海教育出版社，1990：215.

② 中国大百科全书总编辑委员会（教育编辑委员会）．中国大百科全书（教育）［M］．上海：中国大百科全书出版社，1985：202.

③ 中国劳动人事百科全书编写委员会．中国劳动人事百科全书［M］．北京：经济日报出版社，1989.

④ 徐升．人才测评［M］．北京：企业管理出版社，2000：6.

从以上表述可以看出：考试是人类社会各个领域普遍存在的一种社会活动；考试是以人为主、客体的一种对象性活动；考试目的是考试活动形成的前提和运行的归宿；一定的组织形式是考试活动存在与运行的方式；有关资源的选择运用是实现预期考试目的的必备条件；考试的本质是鉴别人的素质水平的个别差异。

## 二、考试的功能

考试自产生以来之所以能蓬勃发展，是由于它具有多种功能。综合中外学者有关考试功能的认识，其主要功能如下。

### （一）考试的社会功能

#### 1. 文化功能

考试已发展成为一种亚文化，是人类文化的重要组成部分。它在人类文化的基础上发育成长，同时又积极作用于人类文化的保存、承传和发展。“从中国文化土壤中产生出来以后，又再创造了中国文化”[①]。若从人类文化演变发展的角度考察，考试通过自身内容的设计，选择、传递、保存、强化现有文化；通过考试结果的使用和人材选拔标准的调控，为社会人才观、教育观、教学观等观念文化的传承与革新导引方向。

#### 2. 行政功能

考试的行政功能，是指考试在提高国家行政治国效能方面所发挥的作用。对于考试在治国行政中的作用，中外历史上的贤哲有诸多精辟论述。早在三国时期，魏人王昶就将年试列为治国的第二大策略，在其《治略五事》中提出：“其二，欲用考试，考试犹准绳也，未有舍准绳而意正曲直，废黜陟而空论能否也”。英国罗伯特·英格尔于1835年预断，中国的人才选拔思想，将来某一天，会像火药和印刷术一样，在欧洲的国家制度中引起另一次伟

① 金诤．科举制度与中国文化［M］．上海：上海人民出版社，1990：1.

大的变革。事实的发展不违所料，科举考试“公开平等、竞争择优”的本质精神，已成为近现代东西方国家甄选治国良才的基本原则，考试则成为控制职业准入，强化劳动行政、人事行政、教育行政的重要措施。由此可见，考试在现代管理中的行政功能是多方面的，它是一种推进各种社会事业的良好法则。

3. 经济功能

考试的经济功能来自于人与外部世界的关系，从根本上讲，考试在伴随人类社会发展过程中，不断地在促进人力资源转化为人力资本，提高人力投资的经济效益，实现人才结构的优化，从而形成最佳群体效应，使人们通过考试活动，能取得一定的外界物，从而满足自己的需要。也就是说，考试活动在人们获得利益和满足需要的过程中具有手段的属性，它是促使劳动力资源转化为能使经济效益成倍增长的人力资本的得力措施之一。

4. 调节功能

考试是作为一种有系统的社会活动参与社会整体运行的，它必须按社会母系统的运行规律调控自身的运行，同时又以其特殊结构所固有的功能作用于相关系统，从而构成社会母系统正常运行的一种协调机制，亦即考试的调节功能。同时，由于考试是社会人才群体系统中的一个子系统，是人才培育系统和人才使用系统的中介，它对人才的培育和使用具有调节与控制的作用。

（二）考试的教育功能

1. 督导功能

考试的督导功能，系指考试对师生选择、追求教学目标的动机，实现目标的意志力及后继人才智能发展的取向等具有强化和导引作用。在具体的教育过程中，教学犹如行驶在知识海洋中的航船。考试的作用，从外部看像航标，从内部看像“调速器”，它导引教学方向，调整教学步伐，强化目标意识，督促师生向既定目标进取。

从教的角度讲，教学过程中的考试，均为检验教学是否实现预定目标而设置。其测试结果总是包含两方面的内容：一是学生掌握知识的数量与质量，技能的准确与熟练程度以及能力的发展水平；二是教师教学的实际成效。因此，考试是检验教师教学是否达到教学目标，促使其自觉将教学目标寓于教学过程，并调整和改进教学内容和方法的有效手段。

从学的角度看，考试是压力更是动力，它能强化学习动机，激发学习兴趣，培养进取精神和促进能力发展，还可以促进学生及时复习功课，巩固和加深知识。同时，科学的考试能使学生加深自我认识，坚定努力掌握科学文化知识的信念和探新求真的志向。

2. 导向功能

考试对后继人才智能的发展还具有强烈的导向功能。此种功能并非人为，而是为考试的社会地位和社会职能所决定，也是考试教育性的集中体现。在学校教育中，考试与教学既有目标上的一致性、又有职能上的独立性。怎样考、考什么、效果如何，都将会立即引起应试者、教师和教育管理人员的高度重视，并使其学习和工作紧紧地围绕着考试开展。考试既反映教学的基本任务，又督促教学双方完成教学任务，并检验教学目标实现的程度，故可能动地作用于教学双方的情绪与节奏，导引教学发展的方向。所以说，考试是一个“指挥棒”。

3. 激励功能

考试作为一种检查学生学习效果的手段有着反馈作用，而反馈结果又对应试者起着激励作用，因为考试结果可以反映应试者的知识掌握程度和能力发展情况以及所存在的问题。当考试成绩反馈给应试者时，就会对应试者产生重要影响，成绩好的学生在感情上表现出一种成功感、优越感、自豪感和满足感；成绩差的学生普遍产生一种羞耻感，主要表现出失落情绪和焦虑心情，他

们害怕别人提及学习成绩，对考试怀有恐惧感。此时，教师应给予正确引导，激发他们的自信心，使他们对考试的反馈信息作出积极的反应，并适宜地调整自己的学习方法，加倍地努力学习。

考试作为一种检查教师教学成果的手段，对教师也有着激励作用。教师从学生的答卷中，从学生的成绩单上可以了解到学生的知识掌握情况，学生的智能发展情况，学生的品行情况等等。好成绩会进一步激发教师提高教学质量的积极性，并激励教师教书育人的热情，使他们迸发出更大的能量，全身心地投入到教学工作中。对于反映在试卷上的、成绩单上和学生思想中的问题，经分析反思后会变为更大的动力，激励他们不断完善教书和育人工作。

考试作为一种检查手段，对教学管理部门的管理工作同样也有着激励作用。因此，考试具有较强的激励功能。

4. 鉴定功能

考试的鉴定功能是通过考试的结果来发挥的。教育管理部门通过对考试结果的分析、认可后，依据有关原则和规定，对应试者、教育者和教学管理人员进行鉴别，以区分优劣、决定奖罚。

其一，对应试者的学业水平进行鉴定。这是考试鉴定功能的最主要的方面。它运用标准分数线来确认应试者的某种资格，是对某种能力、学识、品德、行为水准的认可，还可以区分出名次、等级，以选拔、鉴定出优劣者。

其二，对教育者的教育水平进行鉴定。教师是教育者，他们的教学水平和育人成果主要体现在学生身上，特别是体现在考试试卷上、考试成绩和学生未来的发展上。

其三，对教学管理人员管理水平进行鉴定。管理水平高，就会遵循教育科学和管理科学的规律，使教学能依据教学管理文件，合理地制定教学计划，合理地安排学时，合理地安排理论教学环节和实践教学环节，合理地组织、安排教学；使学生能遵循生理、

心理等规律来最大限度地摄取知识和发展能力，这样学生在考试中就会一帆风顺，取得好成绩。

5. 诊断功能

考试不仅是在考学生，同时也是在考教师，是对教师教学效果的一种检验。教师的教学方法和教学内容，教师传授知识和培养能力的情况，教师的责任心和事业心等等，都将充分地反映在卷面上。由于学生的个性差异，学生自我调整能力不同，需要教师根据考试结果对学生的类型、知识、结构和能力发展水平进行较为合理的诊断，根据诊断结果分清不同的对象，采取不同的方法，矫正错误弥补缺陷。

另外，考试还有利于诊断教学管理部门管理工作的弊端，不断地完善管理。

（三）考试的素质教育功能

考试是素质教育过程的组成部分和重要环节。从教育模式的转变到教育行为的调整，考试改革将在其中起到举足轻重的作用①。

考试毕竟只是测量和评价的工具，因此不能武断地说哪种考试技术服务于素质教育，哪种考试技术服务于“应试教育”。就技术而言，考试技术通用于任何一种教育模式，只是使用的角度不同，任何技术都有自身的特性和缺陷，只是操作者突出使用了哪一方面。因此，需要研究考试与素质教育的统一和它的弊端对素质教育的悖反，使得考试改革的目标和操作都满足素质教育的要求。

1. 考试与素质教育的统一

考试与素质教育从总体上说并不构成矛盾体。素质教育目标是教育活动的出发点，也是教育活动的归宿。考试在教育过程中

① 边星灿．考试与素质教育论［J］．中国考试，1998，（4）：8－12.

按照教育目标的要求，引起、控制和调整受教育者的生理和心理发展。考试过程与素质教育过程之间存在相互制约关系，两者在培养全面发展的人的总目标下构成统一。

人的素质结构是考试命题的基本依据，素质的培养和发展规律是考试测量所必须依循的规律。因此，对人的素质的测量时机、方式、题目以及分数使用方法，不是主观随意决定的，必须满足素质教育的规定性，由此构成了素质教育对考试的制约，取得了考试与素质教育的统一。另一方面，通过考试认识教育的发展形态，测量学生的素质水平，提供改善教育环境，加速素质教育过程的信息，提供促进学生素质水平提高的信息。从这个意义上讲，考试是促进素质教育及受教育者发展的一个动力系统，也是带有强制性的控制渠道，在它的作用下，考试与素质教育再一次达成统一。

体现素质教育的考试改革，在教育公平的基本价值要求上与素质教育的目的是一致的。素质教育思想指导下的校内考试改革，弱化了考试的划分等第和筛选功能，强化了考试的及时分析判断和及时反馈的功能，使考试成为因材施教的调节手段。与“应试教育”最大的不同是，素质教育依据考试信息，调动教育资源向处境不利的学生倾斜，教师无论从情感投入还是个别教学设计方面，都对处于不利地位的学生施以更多的关注。同时，素质教育解决了教育资源过于不平等地被占有的问题，使每个学生至少能够享有基本的资源，使资源分配趋于实际意义的平等，这也是考试与素质教育的统一。[①]

2. 素质与考试是目的与手段的关系

素质的形成是人类文明成果内化为自身较全面的主体性品质

① 胡冬晖，黄崇飞．论考试与素质教育的统一和悖反［J］．成功（教育版），2007，（7）：47－49.

的过程，在这一内化过程中教育无疑是必须借助的条件之一。而作为教育过程中起测量作用的考试，其直接目的是为了提高教学质量；间接目的是培养学生的应试和应变能力；而终极目的是全面提高学生素质。因而，考试就成了一种重要的教育手段，具有工具性，不具有教育的本质属性。作为检查教与学的情况和具体方法的考试，应试教育可以用，素质教育也可以用。应该看到的是，考试本身就具有二重性，既可以激发部分学生的学习积极性，又可以把另一部分学生推向辍学、厌学的边缘；既可以选拔人才，又可能埋没人才。在应试教育思想影响下，考试发生异化，本身从属教育并为教育服务的考试，变成一根凌驾于教学之上的“指挥棒”，应试教育把考试变成了师生共同追求的目标。

与应试教育不同的是，素质教育不是把应付考试作为教学的出发点，而是真正把考试作为教学的一种手段，作为选拔人才的工具，来促进学生全面提高科学文化素质。这就是素质教育与应试教育在考试问题上的分界线。

3. 素质与考试是相辅相成的关系

考试的内容方式是否科学、有效，关键是看其是否真正考出了学生的素质水平。良好的素质一经形成，必然对考试作出积极的反应，使考试取得圆满的结果，并且应当表现出：

第一，高分体现高能。科学的考试，能够全面地考查出每个学生对知识的掌握程度，真正反映学生的素质，高分者绝不是低能者；低能者，或只会记忆知识的人，是低素质的人，必然是考试的失败者。

第二，高能必然经得起高分检验。对高素质的学生而言，考试为他们提供了发挥能力、展露才华的机会，优良的素质是任何考试也考不倒的。

可见，由应试教育向素质教育转轨，并不意味着取消考试，而是关键在于设计什么样的考试和如何使用考试，怎样在素质教

育思想指导下建立促进学生全面提高科学文化素质的考试制度、内容和方法，更好地发挥考试的素质教育功能。

## 第三节　考试与评价

考试是学校教育评价的一种重要方式，也是社会上经常使用的一种选择人才的手段和方式，历来考试与评价都受到社会各界的关注，如何正确认识考试与评价之间的关系，使考试的资源与评价观念和方法更好的相互融合，本节就此问题展开探讨。

### 一、考试与评价的关系

考试与评价是工具和目的、部分与整体的关系，是既有区别又有联系的两个概念：考试主要侧重于测量的技术，侧重于对个体和某个群体的结论性评价，而教育评价更多的是反映教育的价值取向和社会的评价标准；考试偏重于智育，强调考评学生记忆教科书的程度，而教育评价则着眼于学生的全面发展，既要评价学生的知识、技能和能力，又要评价学生的思想品德、兴趣爱好、身体发育状况和心理品质等；考试注重量化，对当前结果作出描述，以区分和选拔学生，而教育评价不仅要对当前结果作出描述，还要考察其发展过程，诊断其症结，提出补救措施。考试为评价提供了分析的基础，没有考试所取得的量化资料，教育评价的定量分析就失去了基础。因此，实施素质教育不仅不应取消考试，而且应强化考试评价机制，加大考试改革的力度，更好地服务于培养人才这个根本目标。

### 二、考试与评价的科学体系

当前，重要的是在传统考试的基础上构建一个科学的教育考试与评价体系，实现从单一考试到多元评价的跃升。

建立科学的考试与评价体系，可以利用考试资源，服务教育评价，即利用评价的观念和方法改造我们的考试。

第一，促使考试具有较多的评价内涵，加强和深化对人的能力、学力和潜质的测评。从考试的角度来说，要加大考试内容和形式改革的力度，使考试既贴近基础教育，同时又有利于高校选拔人才。

第二，利用考试资源做好评价工作，这个评价是对考试的评价、是在一个特定角度对国民素质的评价、是整个教育评价体系的重要组成部分。以考试测量的方式检验和评价基础教育质量是一个世界性趋势。国外不少著名的考试机构参与了一些大型的国际评价项目，如 PISA，TIMSS，NAEP。在这些评价项目中，考试内容不只是来自于教材，更注重来自于现实生活，测量内容不仅是知识，而且是学生的能力。测试中收集的数据信息，不仅涉及智力方面的，还包括认知、态度、情感方面的内容，甚至还有学校、教师在教育教学方面的情况。

总之，利用评价的观念和方法改造考试和利用考试的资源开展评价工作，是实现考试改革和推动教育评价的最佳结合点。目前，建立完善的考试评价体系，发挥考试的评价、导向、咨询功能；建立科学的数据统计的工作机制，收集有效的考试数据信息；建立公平、科学、安全、规范的考试体系，保证学生的全面发展，是促进教育评价，深化教育改革的关键所在，也是时代的迫切需求。

## 第四节　考试与考试学

### 一、考试是一门科学

正确发挥考试的教育性、增大考试的科学性、更加有效地发

挥考试的杠杆作用等一系列问题便成为当今国内外考试研究面临的严峻挑战。正是这些重大的现实问题，要求人们站在马克思主义的基本立场，去创造性寻求解答。同时也说明，考试的研究必须进入理论的领域，才能担负起揭示考试现象后面存在着的客观规律的作用。由此可见，将考试作为一门独立的科学进行研究，既是社会发展的需要，又是考试自身发展的客观要求，同时也是现代科学高度综合而又不断分化的必然趋势。

## 二、考试学的涵义

随着科学技术的迅猛发展，现代物质生产不断赋予考试以新的内容和新的使命，考试便在更为广阔的社会领域得到发展，其体系更加完善，人们对考试的研究和认识也随之深化，愈来愈感到：教育学、心理学、社会学、人才学、管理学、行政学等有关考试的阐述，旨在为其所研究的对象服务，它不能也不可能揭示考试现象的本质规律，满足不了现代社会对考试科学理论的需求。考试因此打破了原居学科的“壁垒”，只身走出了教育等学科的领域，开始了独立成长的生涯。因此，考试学便应运而生。尽管其时日尚浅，但考试学这门新兴科学却已为众多学者所关注和承认。

在我国，从科学的角度研究考试始于20世纪80年代初。最早提出将考试作为一门学科进行研究的是我国著名考试研究专家廖平胜教授。他首次从社会、自然及应用技术科学的角度，以马克思主义的哲学理论为指导，从中国的国情出发，合理继承传统考试理论科学成分、认真总结新中国考试实践的经验、择优吸收国外考试理论研究新成果、系统研讨和揭示考试的内在规律、建立以符合社会发展和现代化考试系统需要的考试理论为原则，构思并形成了《考试学》的基本体系。

那么，什么是考试学呢？

考试学是一门研究考试现象及其规律的新兴社会科学。它以

人类测度、甄别人的素质与智能个别差异的社会活动——考试为研究对象，探求考试发生、发展、演变的成因，研究考试的内在结构、外部关系及其功能，揭示考试的固有规律。考试学的主要任务是论述考试的基本理论、程序、原则与方法，以指导考试实践，并为考试科学的各分支学科奠定理论基础。考试学研究的最终目标，是探索最大限度地开发和利用社会人力资源的有效途径。

### 三、中学考试学的目标内容体系

中学考试学是在《考试学》基本原理的指导下，从中学实际出发，在认真研究基础教育阶段考试和招生考试及其改革的基础上，结合时代特征，从三个方面构思并形成的。

纵的方面，从人类社会发展的角度，阐述考试的发生、发展及与人类社会发展和基础教育的关系。

横的方面，根据不同社会时期的特点和教育规律，论述考试及其改革，以及新课程背景下的学业考试。

现实需要的方面，在大力提倡素质教育的时代，尤其是在深化基础教育课程改革的今天，社会各界对中学的考试，特别是高考的关注越来越多，对科学的考试和公平公正的评价呼声日益强烈。要解决这些问题就必须对中学考试的科学方法进行全面深入的研究。

在内容结构上，分为三个层次。首先，论述考试的基本理论，如考试的本质、原理、意义、历史趋势等。其次，论述科学化考试类别、形式、方法、组织管理与质量统计，评述与中学教育密切相关的“中考”、“会考”、“高考”、“竞赛考试”。第三，论述学生在考试中的心理。

# 第二章　我国考试的历史沿革

在人类文明发展史上，考试是文明的象征。中国是考试的发祥地，早在夏、商、周时期就产生了考试活动，而制度化的考试产生于汉代，是世界上最早实行考试的国家，至今已有两千多年的考试历史。可以说，考试是中国的一大发明。孙中山先生在论及中国古代考试时指出[1]："中国的考试制度，就是世界中最古最好的制度。""至于历代举行考试，拔取真才，更是中国几千年的特色。"中国悠久丰富的考试史是我们值得借鉴和发扬的宝藏，研究中国考试的历史沿革，对我国现行教育考试的改革和发展具有重要意义。

## 第一节　我国古代考试的产生和发展

在我国古代社会漫长的历史中，考试是作为选拔官员的基本制度存在的，基本上分为两种：南北朝以前的推荐制，隋以后的科举制。这一时期，我国的考试从萌芽、产生，到逐步的发展、繁荣，经历了一个相当长的历史过程，不但为我国当时的选士任官提供了有效的选拔手段，为促进社会的发展与稳定、文化的传播发挥了重大的作用，而且为我国近现代考试制度的发展和完善奠定了丰厚的基础，积累了丰富的考试管理经验，至今仍是一笔

---

① 孙中山. 孙中山全集：第五卷 [M]. 北京：中华书局，1985：510－511.

宝贵的财富。其中，科举制在我国古代的考试中占据了绝对的主导地位，其影响之大、范围之广，究竟“利”、“弊”如何，至今仍饱受争议。我们要从历史的角度科学看待科举制，取其精华、去其糟粕，形成对科举考试的正确认识。

## 一、考试制度的起源

考试是文明发展到一定阶段的产物，它的产生是一个渐进的过程。中国的考试究竟缘起何时，学术界众说纷纭。推论最早的是廖平胜所著《论中国考试的起源》一文中将原始社会晚期山顶洞人传授技能时的实践性观察和检验等活动视为“人类考试的胚芽和中国现今考试之源”[1]。也有学者认为，原始社会末期尧舜时代便产生了我国考试制度的萌芽。有史料记载[2]：“试舜五典百官，皆治。”《虞书》中载[3]：四岳举治水，帝曰：“异哉！试可乃已。”而四岳举舜嗣位，帝曰：“我其试哉！”及舜历试诸难，帝曰：“格汝舜，询事考言，乃言底可绩。”这里所说的“试”，不是现代意义的考试，但含有试验、测试的意思。

公元前21世纪，进入文明时代的夏代，职官实行世袭制。由于对外防御、对内镇压反抗与维护社会治安的需要，选拔军事人才成为国家政权中一项重要活动，且纳入国家行政管理的组成部分，史称夏天子“以射造士”，即以训练和比试射箭技能来选拔军事人才。以此，至少可以说作为选拔官吏的考试在我国产生。

另有学者认为，真正具有现代意义的考试制度是在西周出现的。考试制度与国家的产生、学校教育的发展和国家对人才的需求和选拔息息相关。西周是我国奴隶社会发展的鼎盛时期，这时

---

① 廖平胜．论中国考试的起源［J］．华中师范大学学报（哲学社会科学版），1991，(4)：27－31.

② 司马迁．史记·五帝本纪：卷一［M］．北京：中华书局，1972：14.

③ 沈兼士．中国考试制度史（第四版）［M］．台北：台湾商务印书馆，1980.

的国家制度趋于完善，学校也已经盛行和发展起来，“选士”制度成为国家选拔人才的基本方式。“选士”制度是学校选育、乡举里选、诸侯贡士三位一体的模式，采用推荐和考核相结合，由下而上层层推选，有明确的要求、简单的程序及主持考选的各级官员等，是我国考试制度的萌生阶段[①]。

进入春秋战国时代（前770－前221），奴隶制逐渐衰落，新兴地主阶级逐步崛起，“军功”和“养士”成为当时选官任人的主要途径。所谓军功，即奖励战争中有功之人；“养士”就是聚养有才之士，供己选拔任用。此时，周代的世卿世禄制处于崩溃之中，私学盛行，普遍奉行“学而优则仕”的贤才观念，游士入仕游说，各国诸侯竞相延纳。春秋齐桓公元年（公元前685年），任用管仲进行政治改革，抓住人才选拔这一关键，对贡士“设问国家之患”，由国君主持面试（或称口试），这是我国也是世界上口试有文字可考的最早记载[②]。

到了汉代，封建统治者实行“察举”和“征辟”。“察举”是由州、郡等地方官在管辖区内察举人才并推荐给朝廷的一种人才选拔制度。察举制是我国历史上和科举并立，并与之衔接的一种选士制度。汉文帝十五年（公元前165年）九月，“诏诸侯王、公卿、郡守举贤良能直言极谏者，上亲策之，傅纳以言”[③]，这是我国历史上最早的一次策试活动，标志着汉代察举制度的正式确立，也为后来科举制殿试的实行发出了先声[④]。“征辟”则是由皇帝或地方长官直接进行征聘。

另外，汉代太学的课试（即博士弟子考试授官）也是汉代察

---

① 唐群. 科考篇——风云际会考场路［M］. 西安：三秦出版社，1998：192.

② 王兴亚. 孙中山对中国古代考试历史经验的总结［J］. 湖北招生考试，2005，(8)：16－20.

③ 班固. 汉书·文帝纪［M］. 北京：中华书局，1962.

④ 刘海峰. 中国考试发展史［M］. 武汉：华中师范大学出版社，2002：19.

举与考试统一，人才选拔与育人相结合的一条仕进之路。汉武帝创立太学，太学既是当时的最高学府又是国家考试机关。太学的考选制度包括："博士选授"、"博士弟子选试"、"博士弟子课试"等三项内容；考试时间较为固定，实行"一岁皆辄课"（即一年考一次）；考试形式除口试外，也使用笔试，开创了世界笔试的先河；考试方法采用"设科射策"，即事先命多道试题，抄录在竹简上，称为"策题"，皇帝以策题考试称为"策试"，根据试题难度分为甲乙两个档次，由考生随机抽取作答称为"对策"[①]。它既是汉代选士的重要途径之一，也是自成体系的汉代学校考试形式，对我国古代考试的思想、方法、制度等的发展有重要影响。

此外，汉代董仲舒所著《春秋繁录·考功名》一书中最早出现了"考试"一词："考试之法，合其爵禄，并其秩，积其日，陈其实，计功量罪，以多除少，以名定实，先内第之。"意思是对官员进行考核和试用。

因此，有不少学者认为，"最近似于现代考试的，产生于汉代"[②]；汉文帝亲自策问举贤良方正"既是我国不定期举行的特科考试的开端，又是我国取士考试的开端，还是我国考试制度的开端"[③]，标志着我国考试制度的正式确立。

魏晋南北朝时期，由于长期战乱，"人士流移，考详无地"，改行"九品官人法"（即"九品中正"制），辅以察举。"九品中正制"始于曹魏，在曹操"唯才是举"的方针引领下，由各级中正官逐级向上推荐人才。推荐的人才分为九等，前三等为上品，官吏选拔一般从上品中进行。但是，这种选拔制度等于几乎取消

① 高飞．中国古代考试制度简论［J］．教育理论与实践．2004，24（8）：23－24.

② 杨学为．中国需要"科举学"［J］．厦门大学学报（哲学社会科学版），1999，（4）：17－19.

③ 盛奇秀．中国古代考试制度［M］．山东：山东教育出版社，1988：11.

了考试，严格来说并不算是一种考试制度，因此有学者认为[①]：“九品中正制本身并不是考试制度，而且也不是隋唐科举制的前身，它是一种铨选方法而非贡举方法。”

嘉平元年（249年），曹魏征南将军王昶在上呈治国方略五条中，曾精确地揭示了考试的测量功能。他指出[②]：“考试犹准绳也，未有舍准绳而意正曲直，废黜陟而空论能否也。”即考试犹如能衡物体长短的尺度一样，能够客观地衡量人才的知识水平，避免主观空论。这是中国历史上首次对考试作用的准确而形象的论述，在中国考试发展史上具有特别的意义[③]。

学校考试也是魏晋南北朝时期考试发展的重要方面。学校教育已从不分科发展到单科和分科、分馆教学，而且建立了一整套提高教学质量的校内考课法。它与选官入仕紧密相连，与“察举制”共同构成了科举考试制度的渊源。

### 二、考试制度的重大发展——科举制

魏晋南北朝时期的选士制度虽以九品中正制为主导地位，仍然延用了两汉时期的察举制。由于九品中正制存在氏族政治等诸多弊端，以至于出现了“上品无寒门，下品无世族”的腐朽局面，有才华的贫民弟子无法入学入仕，受到了新兴地主阶级的强烈反对。因此，到魏晋南北朝后期，为了削弱门阀氏族的权势，察举制中的考试因素得以加强，考试成为察举制颇具特色的部分，并且扮演了越来越重要的角色，以此孕育了国家设科招考和士人自由报考的分科考试制度的产生。随着社会经济的发展，重视门第名望的九品中正制已不再适应社会的需要，科举制便应运而生了。

---

① 刘海峰．科举考试的教育视角［M］．武汉：湖北教育出版社，1996：14.

② 陈寿．三国志·魏书·王昶传［M］．北京：中华书局，1982.

③ 刘海峰．中国考试发展史［M］．武汉：华中师范大学出版社，2002：28.

### （一）隋唐五代科举制度的萌芽和确立

考试成为选拔人才的主要途径源于科举制的建立。科举制即“开科取士”，就是采取分科考试的办法，选拔封建国家所需要的各种人才的一种制度。

隋朝建立以后，为加强中央集权，于隋文帝时废除了延用下来的九品中正制，以“志行修谨”与“清平干济”二科取士，也就是用德和才两科分科考试，评定录取士人。自此，科举制度开始萌芽。隋炀帝大业三年（公元 607 年），命文武取士五品以上者，按“德行、学业、文才”等十科举人，隔年又增设学业才艺、堪理政事、立性正直、膂力超绝等四科举人，随后在此基础上设置明经进士二科，以试策取士，标志了我国科举考试制度的正式产生。由此，中国的考试完成了从零散的考试现象到成体系的考试制度、从才干的衡量与考试成绩相分离到二者相结合的过渡，形成了一套稳定的考试制度。

科举考试把录取和任用权完全集中在中央，其取士的特点是录取标准专凭试卷，专重资才，而不是由地方察举，实是我国选士制度史上的一次变革和进步。虽然隋代建立的科举制度还处于察举制向科举制的过渡阶段，很不完善，但它开创了中国科举制度的新纪元。

唐朝时，科举制度得到了进一步的发展，科举制度基本确立。“开科取士体制，至唐时大臻完备”。唐代科举的考试形式分为常科和制科两种，还有殿试和武举。常科即每年定期定名目举行的考试科目。制科由皇帝根据需要，临时定立名目下诏考选。考试科目有秀才、明经、进士、明法、明字、明策、道举、童子等。从初唐开设的 90 多科常科、制科中，渐渐独尊进士一科。考试内容以《论语》、《尚书》、《孝经》等儒家经典为主，以孔颖达《五经正义》为标准。考试方法主要为口试和笔试，主要题型有策论、帖经、墨义、诗赋等。考试生源破除门第之弊，推行“怀牒自

列”，无须地方长官察举，到地方政府自愿报名，参加全国统一考试。考试管理由礼部主持，每年二月举行，考场围墙遍插荆棘，外有巡逻军士，考生须经逐个点名，核对举状书，不准携带书籍小抄，一旦入场便不得出入，食宿皆在考场内。但是，唐代科举只是获得出身的考试，出仕任官必须经过吏部的铨选考试[①]。

（二）宋代科举制度的完善和发展

“宋承唐制，抑又甚焉”。宋代在继承唐制基础上，经历了范仲淹、王安石、蔡京三次大的科举考试改革，对科举制有了新的发展与突破，对考试科目、内容和方法等都进行了一系列的改革，克服了创始时期不够完善的缺点，开始向客观化、制度化、严密化迈进了一大步。在宋代，科目分合频繁，但科目设置改革总的趋势是科目日趋简单化，主要有进士科、制科等，武举考试在宋代取得了与文举的同等地位，各科考试的内容和方法也随之不断更改。考试形式上确立了乡试、省试、殿试三年一次三级考试制度。开宝六年，宋太祖亲自批阅试卷，殿试作为制度正式确立。考试管理方面，制定了一系列有助于公平取士的考试规程：“禁公荐”、“停纳卷”；试卷采用“糊名”（即“密封”）和“誊录”；建立锁院制度；实行“别头试”和“锁厅试”。通过制定系统、严密的防范措施，宋代科举考试无论是就制度本身来讲，还是从人才选拔的成效来看，都是历史上最好的时期。其中，“糊名”和“誊录”之法不仅为元、明、清历代科举考试所延用，也被现代考试吸收和借鉴，“密封”至今仍是我们进行考试必不可少的重要环节。在宋代，经过科举出身的人能够直接出仕任官，实现了举和选的结合。至此，科举考试制度发展成熟。

此外，学校教育发展到宋代已日趋完备，“书院教育”和“蒙

① 厉志红．中国古代考试制度的历史演变及特点剖析［J］．河北大学成人教育学院学报，2006，8（4）：66－68.

养教育”在宋代得到大力发展，特别是太学“三舍法”的大力推行，使科举考试与学校考试紧密结合起来，极大地促进了学校教育的发展，进而使科举考试和学校考试在宋代人才选拔与培养中的作用进一步加强。

（三）辽金元时代科举的发展

辽金元是我国三个少数民族——契丹族、女真族和蒙古族掌权的时期。这一时期，各族统治者为了笼络汉族，吸收汉文化的精髓，间或延用了我国传统的考试制度，我国科举考试得到了一定发展。但是由于少数民族当权的特殊性和局限性，科举考试一般仅局限于在汉族内进行，不免带有种族限制，因此这一时期我国考试的发展比较缓慢。

辽代的科举考试是在其“以国制治契丹，以汉制待汉人”的南北分治政策指导下实施的，选拔人才主要是“世选制”。即由帐院推选，一般在契丹族内进行，辽初不设科举保荐之法，后来才逐步向唐宋学习，实施了科举考试，但科场主要面向汉人，专为汉人而设。契丹作为我国第一个少数民族政权，能够接受汉文化，逐步推行科举制，标志着我国古代科举制度的发展已经逾越了民族界限，对加速我国多民族大家庭的发展起到了促进作用。另外，辽代于辽太祖时始建国子监，学校教育得到进一步的发展，这也是契丹人吸收汉文化的具体体现。

金代是中国历史上全面接受和推行科举考试的少数民族政权。金代的科举考试继承辽之制，并采唐宋之法，在女真族特点的基础上，专设“女真进士科”，允许女真人参加考试，但不和汉人混同。还对科举考试进行了一些探索和尝试，如：推行南北选，分南北两场考试，各以经义、词赋取士；实行四级试选制，三年一开科；创设女真进士科，以策论取士。和辽代衰落不振的学校教育相比，金代学校教育在科举兴盛的推动下发展，建立了包括中央官学、地方官学在内的官学教育体系和专门的女真学校教育体

系，学校考试制度也呈现出多样化、民族化的特点，包括国子监学校考试、府州学校考试和女真学校考试，与金代科举取士紧密联系，在当时发挥着重要作用。

元朝建立（1260年）之初，并未实行科举考试，直到元仁宗皇庆二年（公元1313年）才正式制定科举考试的章程。元朝的科举考试体现了明显的民族歧视色彩，把人分为四等，考试形式、内容、放榜方式和任官措施也区别对待。元代的科举考试管理制度已经日臻完善，对考场纪律管理、贡院官员管理等都进行了严格的规定，为明清科考制度的完善奠定了基础。特别是在仁宗皇庆二年（公元1313年）规定将朱熹的《四书集注》作为科举考试的解经标准，从此《四书》与《五经》并列成为我国科举考试的出题范围，被一直延用下来。元代设立的国子学中分斋学习、考试积分制度也汇入了科举考试制度的体系之中，成为科举制度发展的一部分。

（四）明清时代科举制度的繁荣和衰落

明清两朝初建立时，都曾采用荐举和科举相结合的方式选官取士，但是荐举取士的效果不佳，人多而滥，因此科举取士跃而取得了取士制度的主导地位，受到世人推崇。到清朝乾隆时，科举考试的模式基本定型，科举考试得以稳定的施行和延续下去，科举制度也得到了全面的发展。

“科举学校化，学校科举化”。明代借鉴前朝考试的经验，使学校考试与科举考试完全合流，学校考试逐渐沦为科举的附庸，成为了科举考试的预备场所。这一时期，各类学校的考试制度，在原来的基础上和科举考试的影响下，日趋严密和完善，在考试内容和方法上也有所发展。宋朝学校考试中就出现的积分法发展到明朝时形成了比较完备的积分制。如明朝国子监规定，在高年级的孟、仲、季月各考一次，每次试卷分三等，文理俱优的给一分，理优文劣的得半分，文理俱劣的无分，在一年内积分达到八

分的为及格，准予毕业，成绩优异者，经皇帝批准可提前毕业，破格录用[1]。科举考试在考试科目上虽分文武两途，但只设进士一科；在考试内容上使用八股文这种特殊文体；在考试方式上建立了童试、乡试、会试、殿试的固定模式；在取士布局上确立了分地配额的格局。明代考试以科举考试为中心，进一步程式化，达到了我国考试制度发展的最高峰。

清代学校考试和科举考试完全仿行明代的办法，最终使学校考试完全被纳入科举考试的范围之中，建立起了超于前代的完备的考试制度。

明清时期科举考试突出进士一科。考试大致分为以下四个步骤：第一步称为“童试”，也可说是预备性考试。严格地说起来，童试只是一种入学考试，童试合格者只表示已取得了地方官学生员的资格。而真正意义上的科举考试取士制度，是从乡试考举人开始的。

第二步称为“乡试”，是省一级的考试。每三年举行一次，叫“大比”，因为考期定在农历八月，故又称“秋闱”。乡试在各省省城举行。

第三步称为“会试”，是中央级的考试。乡试后的第二年（农历二月）在京城举行。故又称为“春闱”或“礼闱”。

第四步称为“廷试”或“殿试”。会试之后（一般在农历四月）举行，由皇帝亲自主持。殿试只试策问一场，要求考生当场交卷，弥封后送读卷官审阅。殿试并卒淘汰，参加殿试的贡士均能获取进士资格。

明清时期的科举考试制度对学校教育、社会教育乃至家庭教育都带来了深刻的影响。整个社会弥漫着一种劝勉勤学以求功名富贵的思想，形成一种强力磁场，吸引着社会的各个角落和各个

---

① 于信风．考试学引论［M］．沈阳：辽宁人民出版社，1987.

家庭。明清时期科举考试制度越来越严密，规矩也越来越多，防范与惩治办法也越来越精细。尽管科举考试中的防弊措施甚严，但是科举考试中的贪污受贿、弄虚作假的现象仍然十分普遍，诸如贿买、钻营、央带、枪替、割卷、传递、顶名等各种名堂，屡见不鲜。越到后期，科举制度的弊端越加显露。至清时，虽然在继承前代经验的基础上，周密立法，严防弊窦，但千百年来积累起来的科举考试弊端至此发展到了极点。

明清是我国考试制度发展的极盛时期。至此，无论科举考试还是学校考试都已基本定型，形成为固定的规制，在制度上的确实现了规范化、标准化，选拔人才也有了更为客观化、公正性的规则作保证。虽盛极而衰，仍沿用至清末。清末内忧外患，社会变动巨大。中国要开始走上近代化之路，但科举考试却在这种局面前不知所措，改革步伐迟缓，因而严重地阻碍了近代学校教育的发展。最终，科举考试被废除，但作为选拔人才和检测教育效果的考试，却并没有因此而废止。随着西学东渐和欧美列强的侵入，我国的考试也被动地开始了其近代化进程。

从以上叙述中，我们可以发现：我国古代的考试制度与社会发展基本上是同步的。社会越发展，考试制度越完备。凡是社会稳定与繁荣的时代，也是考试制度比较正规与完备的时代；凡是社会动乱的时期，考试制度也不正常。它说明我国的考试制度与我国古代文明的发展之间有着极为重要的内在因果联系。

## 第二节　我国近代考试的演变

从1840年鸦片战争开始到1919年五四运动这段期间，为我国旧民主主义革命阶段，历经清朝晚期、中华民国临时政府时期、北洋军阀时期。为抵抗列强侵略，争取民族独立，这一时期涌现的各改革派开展的洋务运动、资产阶级维新运动、辛亥革命，为

考试改革做出了不懈努力。1919 年以后到 1949 年为我国新民主主义革命阶段，这期间，国民政府对现代考试制度的创立和发展做出了不少有益探索，建立了统一考试制度及毕业会考制度，并将学位制等引入我国。

## 一、晚清科举的停废

在我国传统社会，科举制有其产生和长期存在的必然性与合理性，作为“帝制时代中国最为重要的一项政治及社会制度”[①]，科举在晚清之前的1000 年历史中，曾多次步入行将被废的困境。但是，凭借其强大的历史惯性和修复功能，科举总是与时共进，即罢旋复，以至清代有学者梁章钜认为“终古必无废科目之虞”[②]。然而，进入近代以后，随着社会环境的巨大变化，科举面临西学的强大挑战，虽历经迟缓的自我改造，终抵不住社会发展的滚滚洪流，无法应对内忧外患的困局了。

科举发展到晚清，其弊端日益显露，呼吁科举“补缺改良”的呼声也越来越高。1861 年，洋务运动兴起，洋务派呼吁学习西方的先进技术，革旧纳新，企图用资本主义的手段维护风雨飘摇中的封建统治。封建王朝的典型代表——科举制度也因此受到了不小的冲击，其舆论压力一步一步地变得紧迫起来。在这种渐强渐近的舆论压力下，清廷开始了一系列改革尝试。首先，建立了一批新式学堂，并开设西学课程。1862 年，恭亲王等奏设同文馆并获准设立，1866 年又提议在其中添设算学馆，接着李鸿章、左宗棠、张之洞等又建议设立了各种军事学堂。洋务派仿照西方办学方式，采取新的招考制度，开设西学课程，为我国培养了第一批翻译人员和科技人才。这些新式学堂在旧教育体制上打开了一

---

① 李弘祺．宋代官学教育与科举［M］．台北：台湾联经出版事业公司，1994：14.

② （清）梁章钜．浪迹丛谈：卷五［M］．北京：中华书局，1981.

个缺口，对传统科举考试形成了一个猛烈的冲击[①]。其次，开设算学科考试。1888年戊子乡试开始，单独录取算学科举人，这是中国历史上首次实行中学和西学同考，近代科举改革的努力终于取得了一点突破。但是，算学科考试也只进行了这一次，此后的历科乡试，算学科无不因报考人数过少而改应顺天乡试。

1894年甲午战败后，在举国震惊和指责声中，维新派把矛头直接对准了科举，科举议改进入高潮。1898年进行的戊戌变法对科举的“破”与“革”，主要体现在1898年3月，光绪帝诏令改革武科旧制，停试八股，改试策论。科举考试改革进入到“破”与“立”结合阶段。

1904年1月13日，清廷正式颁布张之洞及管学大臣所订的《奏定学堂章程》，明确宣布十年三科内尽减科举名额。这是废除科举制中决定性的一步，科举制的废除只剩时间问题了。

这一年，爆发在中国领土上的日俄战争及其结果，给科举的废除投下了最后一个砝码。清政府日益暴露的腐败无能以及侵略者的横行肆虐，迫使更多的仁人志士走上了反清的革命道路，孙中山领导的民主革命进入了高潮阶段。这些情势迫使清廷废除科举的步伐骤然加快。1905年9月2日，袁世凯会同张之洞等联名上奏，要求立停科举，以“广学育才，化成民俗，内定国势，外服强邻，转危为安。”清政府迫于内外形势压力，不几日即下令“自丙午（1906年）科为始，所有的乡、会试一律停止，各省岁科考试亦即停止”[②]。至此，运行1300年的中国科举制度正式被废除了。

① 胡向东．民国时期中国考试制度的转型与构建［D］．武汉：华中师范大学，2006.

② 沈桐生．光绪政要：卷31［M］．台北：文海出版社，1985.

## 二、近代学堂考试的兴起

进入近代，学堂就与科举考试紧紧地联系在了一起，与科举的逐步革废相伴，近代学堂考试在新式教育的推动下发展起来，至20世纪初新学制颁行和科举废止，新的学校考试制度终于确立。因此，我国近代考试制度并不是自1840年始，而是起自1862年——中国最早的洋务学堂“京师同文馆”的设立，以及随之而来的学堂考试。

清末学堂大致可以分为洋务派学堂、维新派学堂和资产阶级新式学堂三类。创办时间从前到后，以洋务派学堂最早，从19世纪60年代至90年代；维新派学堂创办于19世纪90年代维新运动时期；而资产阶级新式学堂创办于20世纪初年，以1904年1月13日清廷颁布《奏定学堂章程》为标志，资产阶级新式学堂——自小学至大学在全国各地陆续开办起来。

洋务学堂的考试，引进了一些西方考试方法，弃用科举考试模式，是中国近代学校考试制度变革的开端。洋务学堂考试具有如下特点[①]：其一，学堂考试的目的是选录合格学生和培养外语、科技专门人才，其性质已非选官考试，尽管它有着奖励出身和授予品官的措施。其二，考试内容方面具有双重性，既检测传统知识和基本能力，更主要的是考查西学内容，以西方语言、自然科学知识、军事科学、工程技术等方面知识为主，内容随科目课程设置而定。其三，建立了新的学业考试体系，考试方式多样。一般各洋务学堂都建立了月考、季考、一年一次的升级考和毕业考试这一学业考试体系，具有种类多、考试密集的特点。其四，重视考查专业技术运用能力，具有强烈的实用性特点，开创了近代

① 胡向东．民国时期中国考试制度的转型与构建［D］．武汉：华中师范大学，2006.

实践类操作考试的范例。

20世纪初，清政府开始的所谓"新政"，为学校考试制度的改革创造了前提条件。1902年《钦定学堂章程》（即《壬寅学制》）的制订和1904年《奏定学堂章程》的颁布，标志着近代学校考试已走过1862年以来的萌芽与尝试阶段，开始走向规范化和制度化；而1905年废止科举考试，也客观上促使新式学堂考试转向了全面实施和发展。经过改革，建立了详尽的学制系统，学校考试制度置身于学制系统，学校考试逐步走向严密化和规范化。如将学堂考试分为5种：临时考试、学期考试、学年考试、毕业考试和升学考试；细化考试办法和管理规程；逐步统一入学考试资格；将毕业考试管理权统一归于学校等。这些措施的实施，为民国时期学校考试制度的建立打下了基础。

## 三、民国时期教育考试制度的演变

1911年爆发辛亥革命后，进入民国时期，中国教育在艰难中继续着从传统教育向近代教育的转型，近代学制和教育管理体制得到确立，学校教育和社会教育都得到了一定的发展。与教育的改革和发展相适应，建立了与学制和教育管理体制相匹配的考试制度体系，在考试管理、考试制度构建、考试理论与技术等方面都进行了有益探索。

民国建立之初，即北京政府时期，由于当政者的教育理念、教育发展平衡的水平不高、政局不稳、教育投入不足等因素，教育考试权开始下移给学校，各项考务和考试政策均交由学校管理，国家只做一些原则性规定。这一时期的教育观念向培养知识型、素质型人才转变，否定了以培养官宦为主的封建精英主义教育，体现了从封建主义教育观念向资本主义教育观念的质的飞跃。

但是，考试权的下移也造成了各地各校各行其是、管理混乱、标准不一的问题，引起了诸多关注。基于这个考虑，20世纪30年

代，南京国民政府开始对考试权逐渐上收，开展统一的学校毕业统考制和高等学校招生考试，创立了中学毕业会考制度、学位考试制度和留学考试制度等统一考试的形态，加强政府对统考的管理。这同时也开启了我国教育考试管理权的下移与上收、考试形式的统一与分开之变化的序幕。后来到20世纪40年代，高等学校入学招生考试又进入了多元化招考阶段。

1932年，全国正式推行中小学“毕业会考”制度。同年5月，教育部颁布《中小学学生毕业会考暂行规程》，规定[①]：“为整齐小学、初级中学、高级中学普通科学生毕业程度及增进效率起见，对于所属各中小学应届毕业原经学校考查及格之学生进行会考。”可见，最初的会考是作为中小学毕业水平监测考试进行的，它是国家规定举行、承认结果的省级考试，不仅可以判断学生个人成绩，而且可以判断学校办学质量的优劣，直接影响学校声誉。1933年12月，教育部宣布废除小学毕业会考，并对中学会考制度进行修订和完善。

值得一提的是，作为国家承认的省级学校水平考试，会考在中国考试史上开创了以国家考试行为督导中等学校教育目标实现、整齐中等教育水平的先河，是对我国传统学校考试模式的大胆创新。但由于存在一定的弊端，以及与社会各界的教育理念有一定冲突，因此会考的实施始终伴随着批判的声音。

此外，民国时期，现代学位制和研究生教育也被引入中国。

## 四、留学考试的产生与发展

清末留学考试，是伴随留学教育的兴起而发展起来的。它分为出国留学选拔考试和归国留学毕业生录用考试两类。但出国留

① 杨为学．中国考试史文献集成：第七卷［M］．北京：高等教育出版社，2003：169.

学往往由派遣单位进行选派，并无统一的出国留学考试，而归国留学毕业生录用考试则由教育主管部门统一举办，曾连办7届，考试严密，方法完善，在检测留学毕业生学业后，授予其科名与官职。

民国时期，留学考试获得了新的发展。北洋政府开创留学考核新规章，使公费、自费留学考试初具规范。南京国民政府提高留学资格，强调外语水平，通过举办统一而严格的公费和自费留学考试，较为有效地选拔了一大批优秀人才出国深造，这对于后来我国科学技术和文化教育的发展产生了深远影响。此间，归国留学毕业生考试为新的文官考试所取代，仅举办过特殊的两届留学毕业生甄别考试。

## 第三节　新中国考试改革的历程

1949 年中华人民共和国成立以后，中国考试进入一个新的发展时期。伴随着教育的变革和发展，我国政府和教育部门对各级各类招生考试、学校学业考试等都进行了一系列改革，逐步建立了系统、完备的考试体系和考试制度。

### 一、招生考试改革

招生考试是在各类考试中影响最大，最受关注的考试。在我国政府的努力下，以初中毕业生为主要对象的普通高中和中等专业学校各省统一招生考试制度、以普通高中毕业生为主要对象的全国普通高校本专科招生统一考试制度（以下简称“普通高考”）、以在职在业成人为主要对象的全国成人高等学校招生统一考试制度（以下简称“成人高考”）、继大学本科教育之后的研究生教育招生考试制度（以下简称“研究生招生考试”）相继建立，形成了一个多层次、多规格的招生考试体系。

在这几种主要的招生考试制度中，普通高考作为牵涉面最广、社会影响最大、技术操作和考试运作工程最庞大的一种考试，为其他招生考试提供了原始参照，是各级各类招生考试的焦点所在。自1952年结束各校单独招生，实行全国统一招考至今，普通高考经历了一个创立——颠覆与废除——恢复与改革的曲折过程。1977年恢复高考至今，以全国统一的文化考试为根本的高校招生考试制度20余年不动摇。在这根基牢固、令人倍加珍惜的人才选拔制度背后，改革和发展的脚步却从未停息。

在我国社会经济和高等教育快速发展的趋势下，与普通高考一样，我国的成人高考也经历了从建立到曲折发展，再到不断完善的发展过程，由最初的招生学校单独考试发展到各省教育行政部门组织同类学校进行统一考试，再发展到1986年全国范围内各类成人高校实行统一考试。经过50年的发展，我国的成人高考已经形成现今的统一招考、统一命题、统一评阅等完整的考试制度体系。

相比之下，高中阶段招生考试的改革历程虽不是备受瞩目，显得较为平淡，实际却经历了曲折的发展历程。建国初初中教育普及后，普通高中和初中中专的招生考试竞争日趋激烈。这一时期，我国正需要选拔大量的优秀毕业生进入各层次各领域进行学习、深造，而这些人才的选拔，首先决定于高中阶段招生考试制度的正确与完善。虽然这一时期我国社会主义的教育路线和方针总体是正确的，但是也出现了一些偏差，如五六十年代以家庭出身和表现为能否入学的主要衡量标准；十年动乱时几乎停止了高中阶段教育；拨乱反正后的八十年代，在教育层次中出现了智能倒流，许多成绩优秀的初中毕业生竞相报考中专，仅有一部分优秀的初中毕业生进入普通高中，而更少的优秀高中毕业生升入大学深造。针对这些情况，我国对初中毕业生进行了四股分流，提高了普通高中招生人数，并对高中阶段招生制度进行了相应改革，

确立了各省统一招考的模式。经过几十年的沿革，我国的高中阶段招考制度越来越完善，如今正随着我国教育改革和课程改革的步伐稳步向前发展。

由于我国的中专教育正在改制中，以及研究生招生考试不直接牵涉到中学教育，故未将中专层次的招生考试及研究生招生考试列入本书研究范围。

### 二、学校学业考试改革

学校学业考试是最普通和常见的学校考试形式，区别于会考、高考等省级、国家级考试，它的改革更细致的体现了教育观念的变革和课程教材的实施情况。随着我国基础教育课程改革的一轮轮推行，学校学业考试改革也随之进行了一些变革，如学校学业考试方式改革，学校学业考试内容改革，学校学业考试题型改革等。如今，在新课程广泛实施的背景下，新的课程理念、教学理念以及考试理念正逐步渗透在考试改革中，学校学业考试改革也正在逐步进行。

### 三、初中毕业学业考试改革

初中毕业学业考试是有别于传统观念上的“中考”的一种新型多功能考试。随着新课程改革的深入实施，我国对传统形式上的中考也进行了改革，将初中毕业学业考试作为初中毕业和高中招生制度改革的重要组成部分，既作为衡量学生是否达到初中毕业标准的主要依据，又作为高中阶段招生的重要依据之一。2005年，教育部下发了《教育部关于基础教育课程改革实验区初中毕业考试与普通高中招生制度改革的指导意见》，对初中毕业学业考试科目设置、考试形式、命题要求、成绩呈现，以及初中毕业生综合素质评价制度、高中招生制度进行了一系列改革，改革后的考试制度更符合新课程的要求和素质教育发展的特点。

## 四、高中毕业会考改革

自1986年上海率先实行普通高中毕业会考以来，高中毕业会考已经在我国走过了20个年头。会考实施以来，虽然为我国高中教育的发展起到了一定促进作用，一定程度上体现了其实施的初衷，但在其实践过程中，逐渐暴露了一些弊端，质疑的声音越来越多，要求废止会考的呼声也越来越高。2000年上海市首先废除了会考制度，接着2001年教育部提出将会考权下放至各省区，之后，湖北等省陆续取消了会考，也有一些省份对会考进行改革后仍在施行。会考的存废之争仍在继续，但全国广泛实施新课程后，传统的会考正被新的学业水平监测考试所取代。

# 第三章　考试目标的分类与编制

考试作为一项系统工程，其首要任务就是确定考试目标，即明确考查哪些内容。测验的编制及实施要以预定的考试目标为基础，只有如此考试才能做到有的放矢，避免盲目性。为提高考试目标设计的科学程度，本章将阐述考试目标的分类、考试目标与教学目标之间的关系，以及考试目标如何编制等问题。

## 第一节　考试目标的分类

考试目标就是用教育测量学的术语将对学习成果的评价准确表达出来，使应考者可以了解考试范围和要求。有的学科考试说明中将考试目标称为考试要求，是测试内容的简要说明，是描述、解释考试对象水平和试卷命题的依据，也是评价试卷是否有效的重要依据之一。每 -条考试目标都从一个方面（侧面）描述了考生比较稳定的水平特点。它与考试性质中提到的考试目的不一样，也与考试的具体范围不一样。

### 一、考试目标与教学目标的关系

教学目标是教师希望通过教学活动达到的教学效果，是确定考试目标的唯一依据，缺乏明确的教学目标，就无法确定出测验什么内容、测验的程度如何、以及因何而测试等 -系列问题。基于这 -点，进行测验的第一步就是要把教学目标详细列出来，并

将其陈述为可测量的语言，当教学目标确定并实施测量时，教学目标便成为考试目标。

考试目标就是测验内容，而一个测验所要考查的内容，又应当忠实地反映教学的目标。在教学中，教师是为特定的教学目标而教，学生是为达到这些目标而学，教学目标调节、支配着整个教学活动，考试作为评价、测量教学的基本手段，不能游离于教学目标之外。因此，从理论上说，凡教学目标也都应是考试目标，只是由于一次考试不可能包括所有的教学目标，因而考试目标小于教学目标。但考试目标对教学目标应有充分的代表性是肯定无疑的。

由此可见，进行考试目标研究，必然要以教学目标为本体而展开。而对于教学目标的研究，当前最著名、最有影响的当属美国心理学家和教育家布鲁姆（Bloom，K·S.）和他的同事于20世纪50年代提出的教育目标分类法，教学目标研究基本都派生于此。

## 二、布鲁姆教育目标分类法简介

本杰明·S·布鲁姆是美国教育学家泰勒最杰出的学生，芝加哥大学教育学教授。1956年，布鲁姆立足于教育目标的完整性，制定了教育目标分类系统。他把教育目标分为认知、情感和动作技能三个目标领域，并按照由低到高、由简到繁的顺序把每个目标领域再细分为多个层次和水平，最终完成了《教育目标分类学——认知领域》（1956）这一经典之作。

### （一）认知领域

布鲁姆将认知领域分成六个层次，即知识、理解、应用、分析、综合和评价。这六个层次是按照学生认识的心理过程从简单到复杂、从低层次向高层次依次排列。

1. 知识（Knowledge）

这是最低层次的认知目标。它指对具体事物和普遍原理的回

忆，对方法和过程的回忆，或者对一种模式、结构或框架的回忆。比如对专业术语，对日期、事件、人物、地点，对处理具体事物的方式方法，对各种理论等等的记忆。表述知识目标的行为动词常有“默写”、“复述”、“再认”、“定义”、“指出”、“说出”、“记住”等。

2. 理解（Comprehension）

就是以一种语言或一种交流形式被评述或转化为另一种语言或另一种交流形式时的严谨性和准确性为依据，对材料进行初步的重新整理与排列，指出材料包含的意义和趋势。因此理解包括三个心理过程：转译（Translation）、解释（Interpretation）和推断（Extrapolation）。表述理解目标的行为动词常有“懂得”、“解释”、“翻译”、“说明”、“简单判断”等。

3. 运用（Application）

指在某些特定和具体的情景里使用抽象概念——只是对所学知识的初步运用，而非综合的、创造性的运用。这些抽象概念可能是以一般的观念、程序的规则或概括化的方法等形式表现出来，也可能是哪些必须记住的和能够运用的专门性的原理、观念和理论。表述运用目标的行为动词常有“计算”、“使用”、“修改”、“根据……判断”、“处理……方案”、“对……提出问题”、“把……用于”等。

4. 分析（Analysis）

指将材料分解成各种组成要素或组成成分，弄清各种观念的有关层次，或者明确所表达的各种观念之间的关系，阐明材料各要素的组织原理—即对将材料内容组合起来的组织、系统和结构的分析。表述分析目标的行为动词常有“区分”、“比较”、“整理”、“指出”、“分解”、“确定”、“把……排顺序”、“以……为例写出”等。

5. 综合（Synthesis）

指对各种要素或组成成分加工，形成新的结构整体或新的模式；制定出合理的工作计划或操作步骤；确定一套抽象关系，用以对特定的资料或现象进行分类或解释；或者从一套基本命题或符号表达式中演绎出各种命题或关系。表述综合目标的行为动词常有"分类"、"设计"、"解决"、"撰写"等。

6. 评价（Evaluation）

这是认知领域教育目标的最高层次，是指为特定目的对材料和方法的价值及符合准则的程度做出定量和定性的判断。准则既可以是内在证据，如逻辑上准确性、一致性等，也可以是外部准则，如相应领域的最高标准等。评价是认知能力的最高水平，是认知行为和情感行为的主要结合点和过渡桥梁。表述评价目标的行为动词常有"评价"、"证实"、"审定"、"论述"等。

（二）情感领域

随着认知心理学家对认知和行为、个人经验之间逐渐重视，布鲁姆与其同事克拉斯沃尔（Krathwohl，D·K.）1964年提出将情感领域教育目标由低到高分成五个类别，即接受、反应、赋予价值、组织和形成品格。

1. 接受（Receiving）

包括感觉事物的存在，愿意接受事物，控制或选择事物。接受是情感的起点，旨在培养学生感觉并愿意接受耳濡目染的事件的能力。

2. 反应（Responding）

包括默许反应、愿意反应和满意反应。这是较低程度的积极注意，只是从单纯地感受外界刺激到引起兴趣。

3. 赋予价值（Valuing）

指接纳某事物（比如化学）并认可该事物所赋予的价值，进而爱好相应的有价值之事（研究化学），并逐渐将此爱好稳固

下来。

4. 组织（Organization）

指价值的观念化与体系化，即将各种价值予以整理归类，然后予以系统地组织，使价值间不相冲突，反而彼此隶属，相互作用，从而建立价值概念，组成价值体系。

5. 形成品格（Characterization）

这是情感领域教育目标的最高境界，指价值被内化后，成为学生的价值观或人生观从而约束和统御其言行，价值至此成为品格的一部分。

（三）动作技能领域

布鲁姆早就意识到这一领域的存在，但并没有制定出具体的目标，直至1970年才由布克勒等（Kibler，R·S.，Baker，S·W. & Miles，N·E.）完成，提出动作技能领域的教学目标依次发展的顺序有四个步骤。

1. 整个身体的动作

包括上肢运动，下肢运动，或两个肢体以上的部分肢体的运动。如掷球、跑步、游泳等。

2. 精细的协调动作

包括手与手指协调、手眼协调、手足眼协调、手足眼耳协调。如打字、弹琴、书法等。

3. 非语言交流动作

包括面部表情、手势、身体运动等身体语言。如打手势传递信息、递眼色等。

4. 语言行为

包括发出声音、音与字词的结合，声音投射，声音与手势协调等。如诵读、递眼色等。

布鲁姆的教育目标分类法曾被列为美国教育研究的重大成果而受到广泛重视。世界上不少著名学者的思想，也是根据布鲁姆

的分类思想，再结合教育新理论、新方法或结合各学科实际，加以改造和创新所成。如美国的加涅（Gagne，R·M.）提出的五类教育目标分类（1970），美国的梅里尔（Merril，M·D.）提出的由四个水平十种习得行为所组成的教学目标分类体系（1971）等。

## 三、考试目标的分类

考试目标的分类因标准不同，而分类各异。

### （一）从考试的内容来划分

根据布鲁姆教育目标分类法理论，可以从理论上将考试目标从内容上划分为三个部分：

（1）认知领域考试目标：包括知识的回忆或再认识，以及智能和技能的形成等方面的考核目标。

（2）情感领域考试目标：包括描述兴趣、态度和价值等方面的变化，以及鉴赏和令人满意的顺应的形成。

（3）动作技能领域考试目标：动作技能的考试目标强调肌肉或运动技能对材料和某种操作或需要神经肌肉协调的活动。

我们可将考试目标视为一个大系统，它由以上三个子系统构成。根据子系统的目标要求又可将其分解，形成该系统属下的若干不同层次的亚目标，并最终以可量化的行为动词而体现出来。

在我国，由于受到经济和教育水平等因素的限制，人们对考试目标的内容认识也经历了一个漫长曲折的过程，不同时期的分类标准也有所差别。

20世纪80年代以后，随着我国改革开放和全球经济的不断发展，“知识就是生产力”逐渐成为人们的共识。在教学中，关注知识的系统性和完整性，强调“双基”（基础知识和基本技能），深挖细讲知识点是教师教学的主要内容，所以以知识立意、考察知识点是考试目标的主要内容。

1999年随着教育部颁发了3号文件，要求考试从“知识立意”

向“能力立意”转变，全面考查学生基本素质成为考试目标的重要内容，这时的考试目标主要分成知识目标和能力目标两大类。如化学学科考试目标是对所要考查的化学知识、技能和化学学习能力的描述。从学科角度可以将其归为两个方面：

1. 考试的知识和技能目标

基础知识：主要指普通中学化学课程规定的基本概念和基础理论，常见元素及其重要化合物知识，有机化学基础知识。

基本技能：主要是化学用语，化学计算，化学实验等方面的技能。

2. 考试的能力目标

化学学习能力是与化学基础知识、基本技能联系在一起的。主要是掌握和运用化学知识和技能，如观察能力、实验能力、思维能力和自学能力。由于学科知识结构和学生的认知结构是随着年龄、年级的增长而发展的。所以化学学习能力有阶段性和层次性的区别。根据布鲁姆对认知领域的教育目标分类方法，由低到高分为三个层次，即：了解、理解（掌握）、综合应用。一般认为高层次的要求包含低层次的要求。

了解：对所学化学知识能够正确复述、再现、辨认和直接使用。

理解（掌握）：充分领会所学化学知识的涵义及其适用条件，能够正确判断和说明有关化学现象和问题。不仅“知其然”还能“知其所以然”。

综合应用：在理解所学各部分化学知识的本质区别与内在联系的基础上，运用所掌握的知识进行必要的分析、类推或计算、理解、论证一些具体化学问题。

进入21世纪之后，随着基础教育课程改革的不断深入，人们对教学目标的认识也产生了深刻变化。新课程标准中将教学目标分成三类：“知识与技能”、“过程与方法”、“情感态度价值观”

三个维度，强调每个领域、每个学科的教学目标要注意三个方面的有机结合，整个教学目标的定位从知识中心向能力中心转变，促使学生知识技能学习的过程成为学会学习、形成态度价值观的过程。所以各科考试目标自然分成知识与技能目标、过程与方法目标以及情感态度价值观目标这三种类别。

（二）从考试目标的能力发展水平划分

我国对考试目标的分类研究主要是依据布鲁姆的分类方法，近年来也积极借鉴了加涅等分类体系的研究成果。但即便如此，目前我国还没有一套统一的考试目标分类方法，只是在考试大纲和考试说明中对考试目标做了一些比较含糊的分类，而这些分类基本上都是按照能力水平由低到高的心理发展水平、对考试内容进行重新整合设计而成。根据这种综合式的分类方法，一般可以将考试目标分成：识记、理解、应用、分析和综合、评价、表述这六个层次。如2006年福建省普通高中（历史）会考考试大纲中对考试目标的分类。

A. 识记　再认、再现重要的历史事实、历史概念和历史结论，再认、再现历史阶段特征、基本线索和发展过程。

B. 理解　领会重要历史内容并能作简单的解释和说明，或根据获取的历史知识，进行相应水平的转换、初步的推理，得出正确的结论。

C. 分析　在理解的基础上，能够将重要的历史事实（历史背景、条件、历史现象、历史事件、历史人物活动等）分解成各个方面或各个组成部分，弄清各部分之间的相互关系及其构成的方式，认识其主次、因果关系，从中揭示历史的本质、变化趋势、作用或影响等。

D. 综合　从重要历史事实的不同方面、部分或要素，找出它们之间的内在联系，按照特定要求将历史知识重组、整合。

E. 评价　初步运用辩证唯物主义和历史唯物主义的基本观点，

评价历史事件、历史人物和观点的性质、作用或影响。

F. 文字表达　语言准确，条理清楚。

以下是《中国近代史》（上册）第一章考试目标的具体内容事例：

《第一章　清朝晚期中国开始沦为半殖民地半封建社会》考试目标

A. 识记　危机四伏的清朝封建统治　欧美资本主义国家的迅速崛起和扩张　英国向中国走私鸦片的原因及其危害　林则徐领导的禁烟运动　1840 年英国发动侵略中国的鸦片战争　中国爱国军民的英勇抵抗　中美《望厦条约》和中法《黄埔条约》的主要内容　第二次鸦片战争的爆发　英法联军攻占天津、北京　俄国侵占我国北方一百五十多万平方公里领土　洪秀全创立拜上帝教　金田起义　永安建制　定都天京　太平军的北伐、西征和东征　天京变乱　洪秀全重建领导核心　太平天国后期保卫战　天京陷落　魏源和《海国图志》《康车酋纪行》和《瀛环志略》

B. 理解　《南京条约》及其附件的主要内容　《天津条约》和《北京条约》的主要内容　太平天国运动兴起的原因

C. 分析　两次鸦片战争的原因、性质及其对中国的影响　《天朝田亩制度》的主要内容、性质及其局限性　太平天国运动的性质、特点、历史功绩及其失败原因　林则徐、魏源“向西方学习”的新思想

E. 评价　林则徐　《资政新篇》

由此可以看出，虽然考试目标根据不同标准而分类各异，但基本上都是建立在布鲁姆的教育目标分类学基础上，从不同侧面所进行的新的诠释。

## 第二节　考试目标的编制

作为指导和评价教学的基本依据，考试目标能否发挥其应有的作用，目标本身的设计质量是关键。在进行考试目标的编制时，人们不仅要关注目标的科学性这一基本话题，同时目标的直观性、可操作性等都是决定考试目标的重要因素。

### 一、考试目标制定的依据

我们根据什么规定考试目标，又怎样规定考试的目标呢？其依据有二：一是客观需要。人们根据工作需要提出对考试的期望，期望考试提供什么情况、信息。这是规定考试目标的外部依据；二是考试的功能，考试能够发挥什么功效，能够提供什么样的信息，实现什么样的要求。这是规定考试目标的内部依据。把二者合理地结合起来，使得我们期望于考试的，恰是考试所能实现的；考试所提供的，恰是我们所需要的。这样，就规定出了符合实际（实际需要和实际可能）的考试目标。

《基础教育课程改革纲要》（试行）对中学考试目标有明确规定："国家课程标准是教材编写、教学、评估和考试命题的依据，是国家管理和评价课程的基础。应体现国家对不同阶段的学生在知识与技能、过程与方法、情感态度与价值观等方面的基本要求，规定各门课程的性质、目标、内容框架，提出教学和评价建议"。课程标准中不仅针对各学科的性质和定位做了明确论述，同时针对各科的教学内容和层次要求也作了详细介绍，这为考试目标的制定确立了依据和方向。因此，编制考试目标必须以各学科国家课程标准为根本依据，忠实地反映学生在"知识与技能、过程与方法、情感态度与价值观"三个维度的发展结果。

如前所述，教育目标是考试目标的依据，因此，在确定考试

目标时，除了依据课程标准的要求之外，还应深入研究并弄清课程教学目标，做到测验的目标与教学目标相一致。

但需要指出的是，由于考试的类型和规模不同，考试目标的编制还应依据学生群体的最低限度水准，并紧密结合教材和习题展开。

## 二、考试目标的陈述

布鲁姆认为一个表述得当的教学目标应当具有两个特征：一是必须能详细说明目标内容；二是它应当用特定的术语描述在教学后学生应能做（或生产）以前所不能做的事。考试目标也同样如此。确定考试目标，应尽可能明确、具体，切忌空泛、抽象。空泛、抽象的考试目标不便于理解且不易于掌握。而不便于理解、不易于掌握的考试目标是难于实现的。譬如高考，规定以“选拔最有发展潜力的高中毕业生”为目标，就不如规定“选拔中学基础知识掌握较好、更能适应大学某类专业学习的高中毕业生”为目标易于掌握，便于确定考试内容、分类和选择考试方法。考试目标的明确表述是命制试题和试卷效度检验时具体实施的有力保障，为此，考试目标的表述必须力求界限分明、概括性强、简洁明了、可操作性强。

首先，考试目标中的各个要求之间应该相对独立。虽然从实际情况来看，要做到完全独立非常困难，但是至少应该做到目标之间尽量少交融。如图3－1所示为考试目标之间的关系。




**图3－1　考试目标之间的关系**

甲图中考试目标之间的关系是理想化的，但现实中一般很难达到；丙图中的两者关系最不理想，在实施命题过程中会遇到诸

多疑惑，造成混乱。所以，我们力求要做到乙图的状态，尽量减少考试目标之间的交融，使试题命制和考试目标的确定更加清晰准确。

其次，考试目标的数量不宜过多，也不宜过少，一般10个之内。一次考试能够分辨出考生水平的近10个不同的、稳定的特点，已经足够了，并且也是相当不容易的。

最后，根据考试目标的层次要求采用明确具体、可观察、测量的行为动词描述成行为目标。行为目标是以具体的、可操作的行为形式加以陈述的教学目标，它指明了教学过程结束后学生身上所发生的行为变化，是对学习者通过教学以后将能做什么的一种明确、具体的表述。确定考试目标之后，还必须针对考核内容把它展开成学生"应该做什么"的具体行为目标，这一步就是使考试目标具体化、形象化，变得清晰可见，如使用"写出"、"画出"、"能用"……"计算出"加以描述等。如浙江省高中毕业考试数学考试目标中关于"集合"一节考试目标的描述[①]。

《一、集合》考试目标

1－a　知道集合的概念与符号，记住数集的符号。

1－b　能用列举法和描述法表示集合。

2－a　知道子集、真子集、空集、全集的意义和符号。

3－a　能用交集、并集、补集的意义作出正确的判断，能用图中阴影部分表示交集、并集、补集。

3－b　能写出有限集的所有子集和真子集，能解简单的交集、并集、补集的计算题。

---

① 贾非．考试与教学［M］．吉林：吉林教育出版社，1994，302.

## 三、考试目标的层次

一个考试目标，应该忠实地反映以下信息：其一，考试目标的组成部分，即个体德育、智育、体育、美育、劳动教育五个方面的发展；其二，通过教学形成的个体心理特征，包括知识与技能、过程与方法、情感态度价值观；其三，各部分和各因素的发展水平。（如图 3－2 所示）

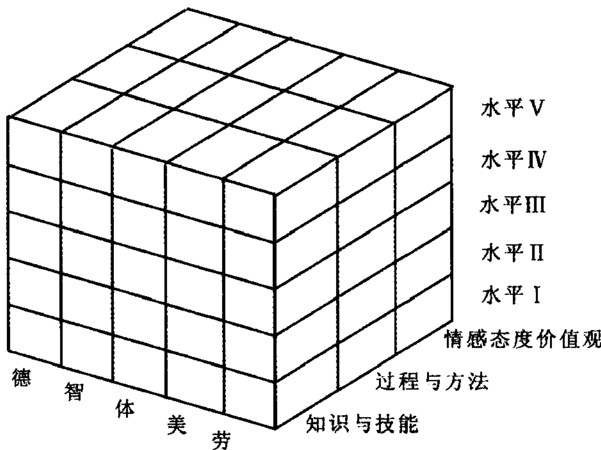



图 3－2　考试目标三维层次图

和教育目标一样，由于知识特点的差异和教学内容本身存在的不同要求，考试目标也存在一定的层次性。我国学者在研究考试理论过程中，受布鲁姆教育目标分类学理论的启发，一般将考试目标分成了识记、理解、应用、综合、评价这五个层次。

最低层次的考试目标一般是识记，主要考察学生对事实性内容的记忆和复述内容，要求能回忆、复述、默写所学的知识。

第二层级的考试目标一般是理解，即能够比较透彻地领会知识和规律，如“能熟练地写出水解的离子方程式并能找出和修正水解离子方程式的错误等”。

第三层次的考试目标是应用，这一层次的考试目标是考查学生用知识解决具体问题的能力。如“能用已学过的函数知识解决一些实际应用题”。

综合层次的考试目标指能在掌握知识的基础上，综合运用有关的知识和方法，灵活地解决一些复杂的学科问题和简单的实际问题。

评价是属于考试目标中最高层次的考查内容，即能应用所学知识，对面临的问题作出科学合理的判断，并预测其可能的结果或问题。

考试目标的水平分类与布鲁姆认知领域教育目标之间，各个层次的对应关系如表 3－1 所示。

**表 3－1　考试目标水平分类与布鲁姆教育目标分类的对应比较**

| 考试目标水平分类 | 布鲁姆教育目标分类 | |
|---|---|---|
| 识记 | 知识 | 学习水平由低到高↓ |
| 理解 | 理解 | |
| 应用 | 应用 | |
| | 分析 | |
| 综合 | 综合 | |
| 评价 | 评价 | |

通过研究实践我们体会到，目前考试目标层次之间的界限尚未达到泾渭分明的程度，相邻的两个目标水平之间有一定程度的交叉、覆盖，或者说界限比较模糊；同时在目标水平的区分上，毫无例外地都依据的是布鲁姆认知领域的分类标准，缺乏全面的层次区分。也正因为此，不同性质的学科，考试目标的层次划分也不尽相同。要使这些界限更为清晰和全面，还有待深入研究。

## 四、考试目标编制步骤

考试目标的编制分为以下几个主要步骤：

第一，根据课程标准和考试类型的基本要求，确定考试的基本内容和范围。

第二，为每个考试内容规定相应的考查水平，如知识、理解、掌握等，并编制成内容目标双向细目表，直观地反应出内容和水平之间的关系。

第三，根据考试内容和考查水平，利用可观察、操作的行为动词确定具体的考试目标。为了能清晰地表达所要阐述的目标，还可采用例证的方式加以补充说明。

综上所述，考试目标的编制技术可归纳为如下程序（图3－3）：

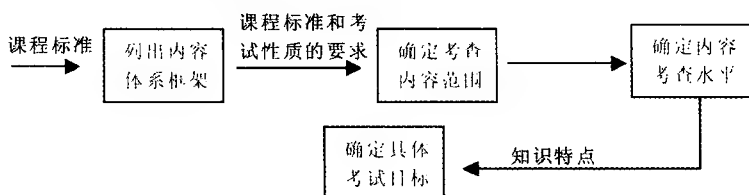



图3－3　考试目标编制程序

当然，不同类型和性质的考试，其目标编制会有所差异，还需要根据客观条件进行调整。如日常学科考试中，除了要依据课程标准之外，还要结合学生实际，只有这样才能体现目标设计的科学性水平。

# 第四章　考试的类型与方法

学业成绩测量的方式有多种，如口头方式、实验方式、作业方式和成绩考试（含测验）方式等。迄今为止，成绩考试仍然是考核学生学业成绩不可取代的主要方式。成绩考试是测量学生在一段时间内对知识的掌握和能力的发展达到何种程度的考试，其目的在于检查和促进教学工作，其内容即是教学内容，其标准即是教学目标。因此，成绩考试也是对实现教学目标程度的测量。

## 第一节　考试的类型

依据不同的划分方式，常用的成绩考试可以分为多种类型。

### 一、按教学进程划分

按教学进程划分，考试可以分为诊断性考试、形成性考试和总结性考试。

（一）诊断性考试

诊断性考试一般是指在学期、学年开始，或一个单元教学开始之前进行的考试。其目的是为查明学生对于即将开始学习的新内容所具有的预备知识和认知能力状况。这种考试可以帮助教师决定施教的基点，有针对性地设计出排除学习障碍的补救措施，也可以为按程度分班教学、个别教学提供信息，因而又可称为配置性考试。这类考试往往在学生无准备情况下进行，所获取的信

息，不宜随意公开评述。

（二）形成性考试

形成性考试是指在教学进程中安排的有计划、系统性的考试。目的在于了解学生对教学内容的掌握程度、教学目标的实现程度，以便为调整和修改课程计划提供信息。如随堂检查性的小测验、单元测验等。这类考试只问及学生是否达到了教学目标规定的水平，不问及差异，惟一依据是教学目标，因而考试的内容应与教学目标相对应，试题应由易到难排列。教师可以从形成性考试中发现问题，及时调整教学进度和方法。学生也可以从中及时发现自己在学习中的缺陷并加以补救。

（三）总结性考试

总结性考试是当教学进入一个学习阶段的末尾，或者进入课程结束阶段安排的考试。其目的是对学生的学习质量和水平达到何种程度作一次全面检查，以便评定和报告成绩，确定升留级，如期末考试、学年考试和毕业考试等。这类考试要求命题与整个教学阶段的教学目标相对应，范围较广而且要有代表性和集合性，试题的难度与区分度必须适中。

## 二、按评价的参照标准划分

按评价的参照标准划分，考试可以分为常模参照性考试和目标参照性考试

（一）常模参照性考试

常模参照性考试的目的在于鉴别或比较被试者之间的个别差异，也就是决定每个被试者的成绩在集体中所处的位置。比较所用的参照点就是考试的常模。考生成绩在常模之上则高于平均水平，在常模之下则低于平均水平。如入学统考、分班考试、高考、学科竞赛一般都具有这类常模参照考试的性质（事实上并未建立正式的常模）。

（二）目标参照性考试

目标参照性考试的目的在于检验被试者是否完成规定的教学目标，以及完成的质量如何。这类考试的参照点乃是事先制订的教学目标或根据教学目标浓缩了的考试目标。评定成绩的方法是把每个学生的成绩与教学目标作比较，而不计较学生个人之间的差异。如期中、期末考试，以及不计入高一级学校招生考试录取总分的毕业会考等都具有目标参照考试的性质。

这两种考试在功能和参照点方面固然明显不同，但这种划分也不是绝对的。常模参照性考试的结果也能反映某一学科部分教学目标的完成情况，而目标参照性考试一般也能反映出被试者在总体中的位置。

目前学校里常采用的考试可归类为表4－1。

**表4－1　学校常见考试的分类**

| 学校考试名称 | 考试性质 | | | | |
|---|---|---|---|---|---|
| | 诊断性 | 形成性 | 总结性 | 常模参照 | 目标参照 |
| 入学统考 | √ | | | √ | |
| 平时测验 | | √ | | | √ |
| 期中考试 | | | √ | | √ |
| 期末考试 | | | √ | | √ |
| 毕业会考 | | | √ | | √ |
| 竞赛性测验 | | | | √ | |

实际上，目前的期中、期末、会考等一般都欠缺事先制订的明确的教学目标作参照，入学统考也并不提供常模，所以以上归类有待探讨。

另外，按试题评分的客观程度划分，有主观性试题考试、客观性试题考试和实验操作考试；按考试目标、试题、成绩是否经过标准化处理，可将考试划分为标准化考试和常规考试。

## 第二节　考试评价的方法

孤立的考试分数是没有意义和教育价值的，只有对考试分数进行正确的解释和合理评价，分数的价值和意义才能充分显示。考试只是评价的工具，如果在考试之后不进行评价，学业考试就毫无意义。

学业成绩评价与学业考试是对应的，有什么样的考试就有什么样的评价。因此，可以按照考试类型进行分类。

依据不同的分类标准，学业考试评价可以分为如下几种类型。

### 一、按收集学业成就信息的方式不同划分

按收集学生学业成就信息并进行评价的方式不同划分，学业考试评价主要有选择性反应评价、论述式评价、表现性评价和交流式评价等。

（一）选择性反应评价

这种方法是指所有的客观试题和纸笔测验的形式。被试者要回答一系列问题，每个问题都有一些可供选择的答案。他们只需从中选出正确的或最佳的那个答案就行了。评价成绩好坏的指标就是答对题的数目或比例。选择性反应类题目包括下面这些形式：选择题、正误判断题、匹配题、填空题。在需要效率时，选择性反应评价是最适用的。但是需要说明的是，选择性反应所谓的客观只是指评分系统，而不能代表题目方面的特点。因为它所使用的题目都是由测验的编制者根据自己的判断和经验来编写的。编制者决定考查哪些内容，编制多少题目，如何表述每一道题，最好或正确的答案是什么，以及在需要时提供什么样的错误答案。因而，无论适用那种方法，所有的评价都来自评价者的主观判断。因此，为了保证选择性反应评价收集信息的准确性，须遵循以下

一些步骤。

1. 制定评价计划

没有任何计划地去编制一个测验，就像没有图纸去盖一所房子。这样会产生两个不良后果：工期拖长或盖好的建筑达不到标准。事先计划周密，测验编制起来就会很轻松。一般有两种计划：测验细目表和教学目标列表。两者选择其一编制计划都很有效。

2. 选取要评价的内容

在制定了测验细目表或教学目标列表之后，就必须围绕要设计的测验题来选取具体的单个知识点和推理形式。即用清晰的语句陈述你希望考查的知识点，语句中要反映出重要的知识内容和答题者必须执行的认知活动。或者希望评价学生应用知识来推断的能力就须在测验中提出认知方面的挑战。

3. 根据命题编制测验题目

在编制题目时须注意几点：清晰地表述重点内容、争取在阅读能力上的要求达到最低、不在题目或题目之间提供正确答案的线索。

（二）论述式评价

论述式评价要求被评价者写出书面的回答。可能是关于具体知识内容的问题，或者对一个复杂问题做出详尽的解答。可能需要比较历史事件、解释理科中的现象、或者解决开放式的数学题，这样学生就必须写出整个的思考和解答过程。测试者根据这些书面论述，然后根据特定的评分规则进行评估。成绩指标一般是解答中的概念性的内容。同选择性反应评价一样，学生的最后得分取决于答对知识点在全部知识点中所占的比例。

（三）表现性评价

在表现性评价中，被评价者实际上是当着评价者的面进行一种特定的活动，评价者观察他们的表现或操作的结果，然后根据他们的表现水平做出判断。表现性评价既可以评价学生展现某种

技能的水平，也可以直接评价学生创作的结果。表现性评价包括两个部分：一个表现性任务或作业，一套评分标准。

评价指标可以是按正确的步骤执行某种操作，也可以是用适当的方式完成某个任务。比如乐器演奏、大声朗读、用第二语言交谈，或者在体育课、舞蹈课上表演某些动作。因此，重要的是“做”。成绩指标通常是一个表现性等级评定或者等级评定的成长记录袋，他们能够反映学生的表现水平。

正是因为表现性评价是在学生执行具体操作时直接观察和评价他们的表现的一种活动，所以主观判断贯穿了表现性评价的各个方面。因此，在制定、实施表现性评价时须注意：

1. 明确评价目的

不管适用哪种评价方法，都必须注意评价的目的。不同的评价目的使用者在不同情况下有不同形式的不同信息需求，他们根据这些不同的信息做出不同的相应决策。

2. 对表现的界定

制定表现性评价的第一个挑战就是如何界定我们所认为的学业成功。这里必须做出两个计划决策：一是选择表现的类型；二是制定评分规则。

3. 设置表现性任务

表现性评价任务就是对学生的挑战。因而，表现性评价任务清晰、明确地反映了学业目标。在设置表现性评价任务时，需考虑两个问题，即任务的具体成分和需要用来反映表现的任务数量。

（四）交流式评价

教师在课堂上收集学生的日常学业信息的一个最普遍的方法就是直接和学生交谈。这种方法通常不能等同于选择题测验或者表现性评价，但直接交流其实是反映学生学业水平的一个好方法。

交流式评价方法包括在课堂上提问和回答、面谈、协商、对话、倾听课堂讨论和口头测验。评价者听取学生的回答，然后或

者根据准则判断对错，或者根据评价的连续体来确定学生处于哪个水平。交流式评价是一种很灵活的评价方式，只要稍作提示就可以进行。通常它包括许多具体的不同形式，如课堂提问、班会和面谈、课堂讨论、口头测验、日志和日记等。

以上几种方法在本质上并没有优劣之分，都不是十全十美，都有其优缺点，并且不可相互替代，在具体实施过程中一定要根据各种方法的特点、适用时间和情形，正确均衡地使用他们。对于具体的一门课程，应采取怎样的考试方法进行学业考试，这需要按照专业和课程的特点，并结合考试的目的、考试的内容、命题的方式等因素做出科学的选择。对于同一门课程，为了全面考核学生的相关知识、能力与素质的需要，可将多种考试方法联合运用，采取分项考核、综合评定的办法。其目的是为了科学、全面地鉴定学生的课程学业水平，促进学生对课程的全面掌握。

最后需要注意的是，评价的作用是为评价者提供反映学生内部认知状态的外部指标。这些指标都是可见的，这样才便于观察、测量或评估。但要想最有效地了解学生的学习程度就必须明确要求学生掌握的那些学业目标，即要根据学业目标选择评价方法。

## 二、按评价所依据的不同标准和解释方法划分

按评价所依据的不同标准和解释方法划分，学业考试评价可分为相对评价和绝对评价。

### （一）相对评价

相对评价的评价标准是建立在参与评价的对象内部，然后把其他对象与标准相比较，以确定其相对位置和优劣程度。如以一个群体（班、校、地区）中的平均状况（平均分）、离散程度（标准差）为常模做评价标准，评价每个学生在此群体中所处的相对位置，以便为区分学生的优劣、显示学生之间的差异和划分等级服务。因此，相对评价又称常模参照评价或相对参照评价。相对

评价不关心学生对教学目标的实现程度如何。

应用相对评价的好处在于确定个人在考生群体中的位置和名次，可以避免教师主观印象的影响，评价比较客观。这种评价又作为编班、升级或择优录取、决定人选的依据，并多用于人数较多的考试。不足之处是只凭卷面成绩作相对比较，而不顾及个人的其他智力因素和非智力因素，且容易引起激烈的竞争，对于学习能力较低者缺乏激励性。

（二）绝对评价

绝对评价的评价标准是参与评价对象之外的客观指标。它是指以预先设定的教学目标为评价标准，来评价每个学生的目标实现程度。因此，绝对评价也称为目标参照评价。绝对评价不关心群体中每个学生所处的位置，只要预先确定达标的标准，将学生的答对率与达标标准相比较，即可做出评价。

试题数量的内容为多少才算达标？要根据考试的性质与目标确定。有的学者提出标准应在 80% ~90%，布卢姆曾以 80% ~85% 为达标标准。一般而言，短期的教学目标要求达标题目可多些，较长期的目标可少些，但至少也要在 70% 以上。

绝对评价一般多用“通过”与“失败”两级标准表示，凡达到指定标准的为通过，否则就被认为未通过或失败。也有用其他标准表示的方法。

绝对评价的优点在于依据教学的各项具体目标选择试题进行测试。其成绩用百分制表示时，就反映了学生掌握相关内容的程度，起到诊断学习存在问题的作用。

（三）两类评价对学生化学测验分数的解释标准与评价方法

1. 相对评价方法

相对评价是以常模参照评价学生的测验分数，先把学生的原始测验分数转换为常模分数，进而指出其在某一群体中的相对位置，以此评价学生学习成绩。这里的参照群体（常模团体）应对

被评价的学生总体有很好的代表性。常模分数就是以常模团体的一般水平（如平均分）作为参照标准的导出分数。在化学学业测量中，最常用的导出分数是百分等级分数和标准分数。

（1）百分等级分数

原始分数仅表示分数多少，不能代表它在全体分数中的地位。而百分数等级分数是一种相对地位量数，它把参加测量的全体人数作为一百份来计算，从而以某一原始分数换算出其在全体中所占的地位，并说明分数比他少的人占总人数的百分之几。例如某一学生一次在化学考试中为 80 分，经换算其百分等级分数是 73 分，即表示参加测试的 73% 的人低于 80 分，并说明超过他的成绩的人占 27%。

计算百分等级的基本公式是：

$$P_R = 100 - \frac{100R - 50}{N}$$

式中$P_R$—百分等级；

R—某分数在全体中按分数高低排列的名次；

N—总人数。

例如　某校长比较高一 3 个班级化学学习成绩的优劣，他先统计了各班的平均分，结果发现平均分相差无几。于是，他从各班各抽取一名得 80 分的学生，并了解该生在班上的名次，然后通过计算对三个班学习化学的情况作了量化的分析，其分析过程见表 4 -2。

**表 4 -2　某校三个班化学学习考试成绩量化分析**

| 班级 | 高一（1） | 高一（2） | 高一（3） |
|---|---|---|---|
| 被抽取学生分数 | 80 | 80 | 80 |
| 被抽取学生名次 | 20 | 18 | 20 |
| 班学生数 | 44 | 38 | 47 |

80分在考试总体中的位置是惟一的，但它在不同班级的位置，则会因人数不同，位置也不同。因此分数的多少与名次的顺序并不能表明考试的确切位置，这就需要用百分等级的方法，计算出每个80分在各班的不同位置，进而看出各班化学水平的差异。这位校长取各班80分学生一名，意在求出各班80分以下学生所占的百分数。由计算公式可得：

高一（1）： $P_R = 100 - \frac{100 \times 20 - 50}{44} = 55.7$

高一（2）： $P_R = 100 - \frac{100 \times 18 - 50}{38} = 53.9$

高一（3）： $P_R = 100 - \frac{100 \times 20 - 50}{47} = 58.5$

由计算 $P_R$ 可知，高一（2）成绩最好，该班低于80分的学生只有53.9%；高一（3）最差，该班低于80分的学生有58.5%；高一（1）居中。

百分等级分数的意义十分明显，计算也较简单，无论测验分数的分布是何种形态，都可以用百分等级表示某学生在团体中的相对位置。但在使用百分等级分数时，要注意以下几点：

①百分等级属于顺序量表，百分等级分数的差异与测验分数上的差异不成比例，其单位是不等的，因而百分等级分数不能进行代数运算，这给对分数的进一步分析带来了困难。

②百分等级只表明学生在团体中的相对位置，并不表明学生掌握所测内容的百分比。百分等级与测验内容答对率不是一回事。

③百分等级对两极原始分数反应迟钝，即处于两极端的原始分数发生较大的变化，也不能引起百分等级的相应变化。如某学生得分为54分是第30名（最后一名），若该生得分为10分，甚至是零分，他仍是第30名。又如测验的第一名（93分）与第二名（89分）相差4分，若第二名提高为92分，他的等级仍是第二名，相应的百分等级不变。显然，两端原始分数的区别在百分等级中

不能灵敏地显示。

④原始分数较集中的地方，原始分数的极小变化也能引起百分等级分数的极大波动。如100名学生是被测总体，其中百分等级从45到55的两名学生的原始分数可能只差3.5分。

（2）标准分数

标准分数有Z标准分和T标准分等，详见第八章。

2. 绝对评价方法

绝对评价方法是以目标参照评价学生的测验分数，即根据预先规定的教学目标，检查学生对教学所要求知识、技能的掌握情况，进而确定测验分数的意义，评价学生的测验成绩。

这种评价不是将被评学生与其他学生或常模团体进行比较，而是与事先规定的、独立于常模之外的教学目标进行比较。学生可以通过目标参照分数，发现自己在目标达成度上的进步，以增强学习动机。常用的目标参照分数主要有以下两种：

（1）掌握分数

掌握分数是说明学生是否达到教学目标所要求的最低标准而设置的一种目标参照分数。在教学中为了了解学生对基础知识和技能是否掌握而设计的测验称为掌握测验。对掌握测验的计分，只需要定出一个可接受的最低标准即可。一般采取“二分法”，如及格－不及格、通过－不通过等等。

掌握分数的确定标准，是根据测验目的和要求的教学实际确定的：在化学教学实际中，常常以60分为及格标准，如果学生的测验分数达到或超过60分，则认为学生达到了教学目标的最低要求。从掌握学习的要求看，一般以75%～85%的正确反应作为及格或通过的最低标准比较合适。

掌握分数常用于那些能够确定出可能接受的最低标准的资格考试或成就考试。

（2）等级分数

在用等级评定的方法表示学生的某种知识、能力、品质等方面的发展水平时，所用的分数就是等级分数。所谓等级分数，就是把要测量的某种能力、技能或品质分为若干个等级，每个等级赋予某一分数。如评定某项实验操作技能，分为5个等级，分别赋予5分、4分、3分、2分、1分。

当常模参照考试与目标紧密相联系时，目标参照分数与常模参照分数之间并无严格的界限。有时掌握分数或等级分数就是同时按照被测团体达到教学目标的平均水平确定的。因此，可以把两种分数结合使用，在确定掌握分数或等级分数时，既要按照教学目标本身的要求，又要依据被测团体对教学目标的平均达成度。上述两类评价的特征及对化学成绩的解释标准见表4-3。

**表4-3　两种评价的比较**

| 类型 | 相对评价（常模参照评价） | 绝对评价（目标参照评价） |
|---|---|---|
| 评价目的 | 考生在考生集体中的位置 | 考生达到预定教学目标的程度 |
| 评价标准 | 考生集体的常模 | 预定的教学目标 |
| 数量表示与解释方法 | 首先将原始分数转换为标准数值，然后才可确定分数的位置。标准数值的表示方法主要有标准分数法、百分等级分数法。如计算出Z分数后通过查正态分布表就可以确定考生的相对位置 | 预先确定达到教学目标的量化标准（达标标准），而后将学生的答对率与标准相比，即可知道目标达成度，做出评价。达标标准有用通过与否的二级标准，也有用三级或五级标准的，一般不可定得过高或过低 |
| 评价功能 | 将考生分等、排队，用于选拔、编班（组） | 衡量达标程度，为改进教学提供反馈信息 |
| 对试题的要求 | 难易适中、区分度好（能把不同程度的考生区分开来） | 注重试题的数量性质与测验内容和范围的一致性，正确反映教学目标的要求 |
| 对应的学校测验 | 分班、编组、升学考试、各类竞赛 | 单元、章节测验、期中、期末、毕业考试 |

| 类型 | 相对评价（常模参照评价） | 绝对评价（目标参照评价） |
| --- | --- | --- |
| 不足之处 | 这种评价方式与教学目标联系不紧密，不能明显看出学生进步程度，可能会引向单纯追求名次分数 | 这种评价的目标分类体系、评价标准、测量与统计等方法和技术尚不成熟。编制能充分体现教学目标的测验题较困难 |

此外，学业考试评价按教学进程划分可分为诊断性评价、形成性评价、总结性评价；按分数解释的参照点划分可分为常模参照评价和目标参照评价。

## 第三节　标准化考试方法

标准化考试是指按照学科系统科学组织的、具有统一标准的、并对误差作了严格控制的考试。它一般包括：命题标准化、施测过程标准化、评分标准化和分数解释标准化四个环节。

### 一、基本程序

标准化考试的程序大体可按图 4－1 进行。

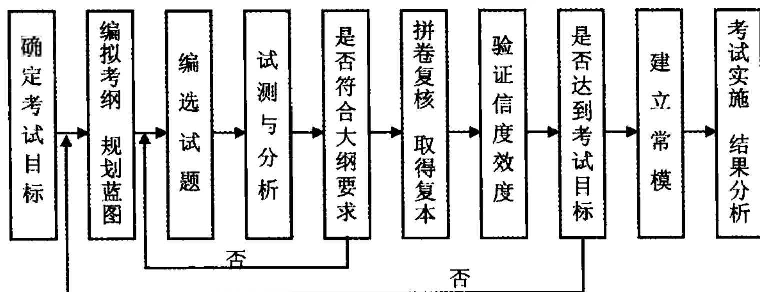



图 4－1　标准化考试基本程序

在上述基本程序中，以下几个步骤最为重要：

（一）编制考试大纲和设计组卷蓝图

考试大纲要明确考试的目标、范围、内容、题型、答题方式、时间要求等，并附有模拟题。一般提前半年左右公布，使教师和考生心中有数。

组卷蓝图即编题计划，其基本形式是含有考试目标和考试内容的命题双向细目表。它指出了应该编写多少和编写哪些种类的题目，以及每种内容的分数分配。命题双向细目表是考试大纲的具体体现和命题的依据。

（二）试测和拼卷

试测是标准化考试编制程序中的关键，其中心环节是项目分析，即对编选的测题结果作统计分析，以确定测题的难度、区分度。标准化考试要求题目的难度在0.3～0.8之间，区分度在0.3以上。试测合格的题目不能当年使用，可以制成题卡存入题库备用。

拼卷可以从题库或试测题中选择难度适宜、区分度高的题目去组卷，并编出多份等值复本。

（三）验证信度、效度，建立为解释分数用的常模

目前标准化考试多为常模参照性考试。标准化考试一般要求信度在0.9左右、效度在0.4～0.7之间。达到要求了就建立常模，这样才能实现分数解释的标准化。

（四）考试的实施

考试实验必须在严密的组织程序和等同的施测条件下进行，一般编有《考试指导手册》供主考者使用。对试卷的印刷、保存，考试的组织，阅卷评分等方面均应实现标准化。

## 二、主要优缺点

（一）优点

1. 试题编制严密，知识覆盖面广。

2. 经过试测，反复修订，有确定的考试质量指标。

3. 施测过程和记分严格、统一、规范，评分客观，阅卷迅速。

4. 建立了常模，使分数有可比性，并可用来校准其他考试。

5. 有利于化学教学，避免盲目地搞题海练习。

（二）缺点

1. 编制复杂，教师难以自编标准化测验。

2. 在测量高层次的学习水平方面不如常规考试方法那样灵活方便。

3. 大量使用选择题，有时出现猜测得分的误差，因此对选择题的编制要求也很高。

4. 多用于常模参照测验，特别强调常模，因而若处理不当，易于引起追求名次，忽视教育应有目标的追求。

鉴于上述各点，在常规教学的考试中，更应提倡模仿标准化考试的做法，使教师自编测验在命题形式和施测过程方面趋向于科学化、标准化。

## 三、澄清几点误识

（一）标准化考试概念的理解

1. 不能说以选择题为主的考试就是标准化考试。虽然标准化考试大多采用选择题等客观记分的题目，但二者绝非等同。

2. 用计算机阅卷并不等于标准化考试。因为全国高考的部分试题就是用计算机阅卷，但不能据此说这就是标准化考试，或说其中标准化考试部分的题目用计算机阅卷。

3. 用标准分记分也不等于标准化考试。不能简单地用是否以标准分记分作为标准化考试的标准。

4. 标准化考试不仅仅包括常模参照性考试，也包括目标参照考试。因此，必须建立用做比较解释的分数标准“常模”或“目标”。

5. 某些部门不经过试测，也未建立常模或目标，就把一些粗

制滥造的选择题拼成汇编，美其名曰《高考标准化试题》，这是对标准化考试的曲解。

6. 标准化考试一般由专门机构编制，关键在于试题的编制和检测，而不是以考试规模论定是否为标准化考试。

（二）常模与试测、正式考试的关系

1. 常模是通过标准化考试样本在试测中、正式考试前得到的。它的功用是为以后的等值考试作评价参照。

2. 常模也可以在正式标准化考试中建立。因此，有的书把建立常模作为标准化考试的程序的最后一道工序。但这只能作为今后等值性考试的参照标准，实际上这个正式考试对今后等值性考试而言仍属“试测”。

3. 正式考试得到的常模与试卷设计时依据的常模（试测所得）是基本一致的，这是由试测带来的。

4. 常模一旦建立，一般可稳定多年，因而以后数年都可按此编制等值性试卷，而不必要再一次次试测求常模。

## 四、广东省1987年首次化学标准化高考简介

（一）制定考试大纲

其根据是中学化学教学大纲、教材及高考目标、要求，判定的具体内容主要包括：

1. 考试目标和要求；
2. 考试范围和内容；
3. 考试方式和试卷安排；
4. 评分记分的方法。

其中考试目标确定为：考查学生对中学化学知识学习情况，鉴别考生的知识水平和技能。具体分五个层次：识记、理解、应用、分析和综合、探究。详见第三章第一节。

考试范围和内容确定为五个部分：基本概念和基本理论、元

素及其化合物、有机化合物、化学计算、化学实验。大纲中在各部分之下又详列了具体的知识内容，分别是七、九、四、八、六大项目。

考试方式为书面答卷形式。试卷分为第一卷和第二卷，第一卷由阅卷机、电子计算机评分统计，第二卷先由教师手工评分记分，再用计算机统计分析。

考试大纲制定后，由高考主管部门审批，并提前半年向全体考生公布。

（二）编制“命题双向细目表”

依据考试大纲编制的考试目标、考试内容的双向细目表中，用数字具体地规定了各目标与内容之间的比例，如考试目标在“识记”方面占6%，其中基本概念占1%，元素化合物占3%，有机物占1%，化学实验占1%，见表4－4。

**表4－4　化学高考标准化命题双向细目表**

| 考试内容 | 考试目标和分数分配 | | | | | |
|---|---|---|---|---|---|---|
| | 识记 | 理解 | 应用 | 分析综合 | 探究 | 总分 |
| 基本概念基本理论 | 1 | 13 | 9 | 7 | 2 | 32 |
| 元素化合物 | 3 | 5 | 6 | 5 | 2 | 21 |
| 有机化合物 | 1 | 5 | 3 | 4 | 2 | 15 |
| 化学计算 | | 3 | 4 | 8 | | 15 |
| 化学实验 | 1 | 6 | 2 | 6 | 2 | 17 |
| 总分 | 6 | 32 | 24 | 30 | 8 | 100 |

（三）着手建立题库

具体方法见第六章第三节。

（四）命题与测试

从1986年6月至1987年3月，共编制了三次测试题，在全省试测，淘汰了难度低，区分度差（0.20以下）的试题，编者期望

考试的信度在0.80以上，大多数题的难度在0.3－0.8之间，整个试卷的平均难度在0.6左右，试题的区分度不低于0.30，考试得分的标准差在15左右。

（五）建立解释分数用的常模

经二次试测和四次统计分析建立了常模。其期望值是平均分60分，标准差14～20，作为编高考卷依据，在高考后是将原始分数转换为T分数向考生公布成绩的，转换公式为：

$$T = 14Z + 60$$

（六）考试实施

略。

（七）试卷分析

分析结果表明，第一卷难度适中（0.56），区分性能好（0.314），信度较高（0.914）。第二卷难度稍偏易（0.64），但区分性能比第一卷还好（0.367），在题量较少的情况下信度达0.75还是可以的。整卷难度适中，区分性能良好。这表明，我国化学高考首次进行标准化考试的质量是良好的。

## 五、我国标准化考试的展望

我国从20世纪80年代中期，在广东省等地开展了标准化考试的试验和研究工作。到1993年，广东省的标准化考试试验告一段落。目前，标准化考试的理论与实践已在教学领域里产生了深刻的影响，其发展趋势主要表现为：

1. 全国高考命题正逐渐朝着标准化方向发展。可以说，近年实行的是模拟标准化高考。

2. 我国高考将继续吸收标准化考试的优点，克服其不足，坚持主客观题型（选答与供答题型）相结合这个具有我国特色的方向。

3. 标准化考试的思想和程序日益渗透到教学之中，常规考试

中的教师自编测验已日趋科学化、标准化。

## 第四节　常规考试的科学化方法

常规考试是课堂教学中由来已久的最常用的测量学业成绩的方法，故又可称之为传统考试方法或教师自编测验方法。其特点是以主观性试题为主，题量较少，题分较大，偏重于考查学生的理解、逻辑思维能力和文字表达能力。其优点是不仅可以逐点考查，还可以综合考查，因而可以系统检查学生的理解能力、思维能力、表达能力，测量较高能力层次的学习水平。并且命题用时少，容易实行，答题时猜测得分机率低，因此一直是现行学校考试中最常用的方法。常规考试主要适用于提供教学反馈信息，考核学生综合运用能力和考生数量较少的课堂考试。

但是，由于这种考试的命题主要依靠教师的个人经验和水平，因而具有命题不稳定、知识覆盖面小、评分欠客观、分数的解释与合成不甚合理等不足。

### 一、科学程序

针对常规考试方法的不足，可以借鉴标准化考试的理论与方法，对其加以改进。我国高考在这方面创造的成功经验就值得借鉴。科学、规范的现代常规化学考试一般包括三大要素：考试的设计、考试的组织实施、考试成绩的使用。每

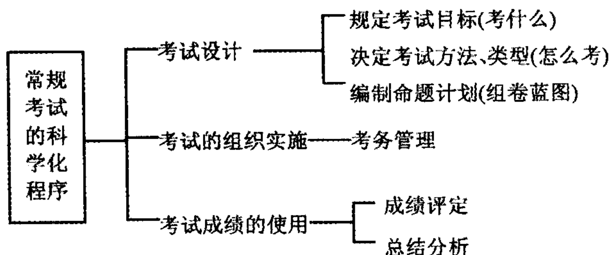



图4－2　常规考试的科学化程序

个要素都应有一套科学严密的程序，见图4-2。

## 二、操作过程

### （一）确定考试目标

即根据教学目标确定要考什么。

### （二）编制命题双向细目表

命题双向细目表是试题试卷设计的“蓝图”，是命题的依据。它从考试的知识内容、测量目标两个维度出发，对各类知识应达到的水平、分数的分配、试题的数量等做出规划。教师自编测验的双向细目表，可以是一份集上述内容为一体的总表，也可以分解为下述分表。

1. 知识点和教学目标双向细目表

此表可参照执行教学大纲、考试大纲制订的教学目标，并在平常的教学工作中逐渐完善。如先编出教材各章节和单元的知识点、教学目标双向细目表，在此基础上编出各年级总的知能双向细目表。

2. 试卷分数分配和知能双向细目表

这是按应达到的教学目标来分配分数的双向细目表，表中的分数也具体规定了各目标与内容之间的比例，如表4-5。

**表4-5　高中化学毕业会考知能分数分配表**

| 考试内容 | 考试目标和分数分配 | | | | | |
|---|---|---|---|---|---|---|
| | 知识 | 领会 | 应用 | 分析 | 综合 | 分数% |
| 基本概念 | 4 | 4 | 2 | | | 10 |
| 基础理论 | 4 | 10 | 6 | 4 | | 24 |
| 元素化合物知识 | 4 | 4 | 8 | 7 | 5 | 28 |
| 有机化合物 | 4 | 6 | 6 | | | 16 |
| 化学计算 | | 2 | 8 | 4 | | 14 |
| 化学实验 | 4 | 4 | | | | 8 |
| 分数% | 20 | 30 | 30 | 15 | 5 | 100 |

分数分配的比例要根据考试性质而定。高中毕业会考是属于目标参照性质的考试，以检查达到目标为主，则知识、领会、应用与分析综合可分别占80%与20%左右。若是升学、竞赛之类常模参照性质的考试，提高目标可占多些，两者可分占60%和40%左右。

编制试卷知能分数分配表的主要依据是：教学大纲和教材各类知识的容量，教材中习题的水平分布等。如中学化学教材以基本概念、基本理论和元素化合物知识为主，习题的能力水平以领会和应用为主，因此，分数分配就应在这些方面的比例大些。还可采取请专家用投票的方法确定。

3. 试题配方双向细目表

此表是在前述工作的基础上拟定的，其目的是为了避免内容的重复，为试题的挑选指明方向。表中具体规定了某次考核的知识点、目标和题号、题数、题分。故又称之为命题计划书。此表应具有保密性。

（三）编选试题和组卷

为了自编测验的选题方便，教师可以在平时的教学中建立自己的“小题库”。如每次测验后统计一下各题的难度，而后制作一个小卡片，将其按知识点、学习水平编号，抄入题的内容、答案，写上题型及难度值。有条件的也可求得区分度。

（四）制订考场组织制度和考试纪律

（五）准备标准答案和规定记分规则

（六）科学解释分数

由于测验的原始分数最大的缺陷是不具有可比性，由此不能孤立地对其加以解释和比较。对于不同目的的测验，也要用不同方式来解释分数。平时测验多属于目标参照性测验，应以教学目标为评价依据，用达成度来解释分数（达成度＝实得分/应得分。应得分即分数分配表中各项分值）；其结果不可用做排列名次。即

使两人成绩都是80分，但因失分情况不同也不能简单地说其学习水平相同。对于升学考试、竞赛之类具有常模参照性质的测验，教师应学会用统计处理的方法，将学生的原始分数转换成标准分再作评价。

## 三、自编目标参照性测验示例

### （一）高中有机化学测验试卷知能分数分配双向细目表

表4-6　高中有机化学测验试卷知能分数分配双向细目表

| 知识内容 | 学习水平和分数值 | | | | | |
|---|---|---|---|---|---|---|
| | 知识 | 领会 | 应用 | 分析 | 综合 | 分数% |
| 烃 | 5 | 15 | 10 | | | 30 |
| 烃的衍生物 | 5 | 20 | 10 | 15 | 10 | 60 |
| 糖类蛋白质 | 5 | 5 | | | | 10 |
| 分数% | 15 | 40 | 20 | 15 | 10 | 100 |

### （二）试卷

1. 选择题（每小题均有1个或2个正确答案）（共70分）：以下共28小题，每小题2.5分。题略。

2. 完成以下指定的化学方程式（有机物一律写结构简式）（共5分）：共2小题，题略。

3. 判断以下物质（共8分）：题略。

4. 计算（共7分）：题略。

5. 写出由乙醇转化成甲醇的各化学方程式（有机物一律写结构简式）（共6分）。

6. 设计一个简单的方案，除去以下有机物中的杂质并简述实验的过程（要求有加入试剂、主要操作、主要仪器）（共4分）：共2小题，限不超过25个字。题略。

此试卷内容与教学目标一一对应，通过学生对试题内容作答

的情况，可以方便地评价其教学目标达成的程度。

（三）目标达成度表

**表4－7　某学生完成以上试卷的目标达成度表**

| 知识内容 | 学习水平和得分（实得分/应得分） | | | | | | |
|---|---|---|---|---|---|---|---|
| | 知识 | 领会 | 应用 | 分析 | 综合 | 实得总分/应得总分 | 达成度（%） |
| 烃 | 5/5 | 10/15 | 5/10 | | | 20/30 | 66.6 |
| 烃的衍生物 | 5/5 | 20/20 | 5/10 | 5/15 | 2/10 | 37/60 | 62.0 |
| 糖类蛋白质 | 5/5 | 5/5 | | | | 10/10 | 100.0 |
| 实得总分/应得总分 | 15/15 | 34/40 | 10/20 | 5/15 | 2/10 | 67/100 | 67.0 |
| 达成度（%） | 100 | 87.5 | 50 | 33.3 | 20 | 67 | |

# 第五章　考试的题型与编制

任何试卷都是由一定量的试题组成的，由于试题的类型不同，因此在考查功能、编制技术和命题规则上也存在着一定的差异。

试题是考试中最小最完整的测验单位，是体现考试目标内容的载体。任何试题都包括四个要素。

（1）题干：提供问题的情景和条件。

（2）题问：提供问题的方向、范围和性质。

（3）题解：提供解题的方案。

（4）题规：提供实际测验和评分的规则。

其中，题干和题问在卷面上，而题解和题规放在答案上（选择题例外）。各种考试能否准确地考查出要考查的行为特征，同选用试题的题型和质量有密切的关系，所以我们有必要研究试题的题型和编制要领。

## 第一节　试题的类型

根据不同的标准，试题可以分成许多不同的种类，最常见的分类有两种：

### 一、按评分方式分

#### （一）主观性试题

需要评卷者作主观性判断评分的试题。如：问答题、实验设

计题、小论文等。

（二）客观性试题

答案确定、评分客观的试题。如：选择题、填空题、是非题等。

## 二、按考生作答方式分

可分为选答题和供答题两大类。

（一）选答型试题

选答型试题不仅向考生提供题干、题问，而且向考生提供题解，但正确的题解或最理想的题解需要考生对题干、题问作出反应后进行选择。选答型试题又可以分为以下几种：

1. 二项选择题

二项选择题又称是非题或正误题。通常是给出一个句子，要求考生对叙述的信息作出正确或错误的判断，主要功能是考核学生对重要的概念、原理或事实方面理解的确切程度和容易混淆的相似材料的辨别能力。

这类题目可以在短时间内考核多个知识点，具有编撰较容易、评分标准客观单一、不受主观因素影响、书写方便迅速等特点，但是随机猜测概率高达50%，所以真实性较差，因此只限于临时的形成性测验。

对二项选择题还可以进一步分类，如按对考生的要求划分，可分为真伪式、正误式、订正式、是非式；从回答问题所需的能力上划分，可分为测试记忆水平的是非题与理解水平的是非题。

[范例] 下列说法是否正确？

①在原电池中，活泼金属是负极。　（　　）

——记忆水平

②将0.1mol·$L^{-1}$的醋酸溶液用水稀释10倍，醋酸的电离度

增大，这个溶液的氢离子浓度也同时增大。　　　　　　　　（　　）

——理解水平

2. 多项选择题

这类试题一般由一个题干和若干个备选项组成，题干提出问题，选项中有正确答案（正道）和迷惑答案（误道），要求选一个或不止一个正确或最佳选项。这是一类最常见的选择题，其应用非常广泛，可以考查学生的基础知识和基本技能，考查分析、比较、判断、推理能力。由于多项选择题还可以在题干和题问上进行变化，题干和题问可以是有关该学科知识的直接提问和不完全表述，也可以是因果关系的陈述句、图表、分类等内容，所以又可把这类选择题进一步分成单解选择题、最佳选择题、识图选择题、分类选择题、多解选择题等多种形式。

（1）单解选择题：单解选择题是选择题中最广泛的一种题型，题目指出诸多选项中符合题意的选项只有一个。单解选择题如果能确认正确选项，其他选项可不予分析。测试的内容是单一结果时，只能采用这种题型。

[范例] 元素X有质量数为16和18的两种同位素。X元素的原子量是16.4，则同位素$^{18}X$在自然界占的百分比是　　　　（　　）

（A）80%　　（B）75%　　（C）20%　　（D）50%

单解选择题常用来测试考生的快捷反应能力。这就需要考生不仅要牢记重要物质的基本常识，对概念理解深透，而且要有敏锐的洞察力。

由于单解选择题明确指出正确答案只有一个，具有一定的猜测性，如四选一题答案猜中率为25%，故多在对难度要求较低的考核中采用此类题型。

（2）最佳选择题：这类试题的特点是供选答案中有一个答案明显优于其他答案，正确选项只有一个。

[范例]《祝福》揭示祥林嫂之死的最根本的原因是　（　　）

（A）以柳妈为代表的封建神权。

（B）以鲁四为代表的封建政权。

（C）以封建礼教为核心的整个封建宗法思想。

（D）以婆婆和大伯为代表的封建夫权和族权。

最佳选择题常常是在要求考生从几个可行的备选答案中，选择最好的或最符合题意的答案。它可用来测试考生的分析、比较、评价等能力。这类选择题的迷惑答案更具有诱惑性，因此，具有一定的难度。解答这类选择题时，一般不宜用直觉去判断，而应逐一分析各选项，全面考虑，对照比较，确定最佳答案。

（3）多解选择题：多解选择题有两个或两个以上的正确选项。它只有正确，错误之分，没有最优、最差之别。题干不能提“最……”要求；也不能用“以上都”作为答案。它的各个选项都应围绕着同一中心，否则就会变形为几道不相关的是非题。这种题型多用在一个问题有两个以上答案，或一个问题需从多方面考虑的情况。

[范例]配制与保存硫酸亚铁溶液，应当采取的必要措施是　（　　）

（A）把蒸馏水煮沸，以赶走水中溶解的氧气

（B）溶解时加入少量稀硫酸

（C）加入少量还原铁粉并充分振荡

（D）加入少量盐酸，并转入棕色瓶里

多解选择题可以测量考生各种层次的学习水平，尤其是在测试考生对知识的全面掌握程度以及从多角度、多方位综合思辨的能力更有其独到之处。所以在对难度要求较高的考核中大多使用多解选择题。

（4）因果选择题：这类选择题的题干由陈述句和理由句两部分组成，依据这两部分的关系，选择约定的字母做答案。如下表：

| | A | B | C | D | E |
|---|---|---|---|---|---|
| 陈述句 | 正确 | 正确 | 正确 | 错误 | 错误 |
| 理由句 | 正确 | 正确 | 错误 | 正确 | 错误 |
| 二者关系 | 有因果关系 | 无因果关系 | / | / | / |

［范例］有机物燃烧时，一定都生成二氧化碳气体，因为有机物都含碳元素。（　）

因果选择题不仅可以测试考生对基本概念、基本理论的记忆与理解，测试考生能否应用有关概念、理论分析说明一些学科实际问题，能否从一些具体事例中正确抽象、概括出有关的概念，还可以测量学生在该学科中的理解、应用水平，所以难度较高。正确解答这类试题时要求学生对概念要明白，搞清楚“应该怎么说”，“为什么要这么说”。解题过程应先把结论和理由分开，再对二者分别分析研究，最后再考虑二者的逻辑关系。

（5）比较选择题：比较选择题具有鲜明的结构性，一般有四个供选答案。四个选项的基本模式是：只是 A；只是 B；既是 A 又是 B；既不是 A 又不是 B。

［范例］甲酸分子具有羧基，又具有醛基，甲酸应具有（　）

（A）羧基的酸性　（B）醛基的不饱和性

（C）既具有 A 又具有 B　（D）既不具有 A，也不具有 B

这类选择题不仅可以考查考生对知识掌握的全面程度，也可以测量考生对两类以上情况进行比较、鉴别的能力。

（6）分类选择题：这类选择题一般有两种形式，一种是“找出一组事物与众不同的一种”，另一种是“找出一组事物归类在一起的理由”，然后从供选答案中找出跟它们属同类的事物。由于这类试题需要在分析的基础上实施，所以可以测量综合级水平。

[范例] 选出加点字在意义上不同于其他三个的一项　（　　）

（A）危如累卵　　（B）危言耸听

（C）危在旦夕　　（D）转危为安

（7）排列选择题：这类试题通常由一组按一定的规则排列的数字或物质组成。在排列的顺序中常有一个空位，要求考生找出排列规则，再从备选答案中选适宜的数字或物质填入空位之中。

[范例] 排列 $NH_3$、$N_2$、（　　）、NO、$NO_2$，请根据其排列规律，从下列物质中挑选适宜的物质，并将其序号填入括号中。

（A）$N_2O$　　（B）$NaNO_2$　　（C）$HNO_3$

（D）$NH_4NO_3$　　（E）$N_2O_4$

排列选择题可以测试考生观察分析综合等能力，也可以测试考生对知识的纵向联系掌握的水平，但是编制较为困难，使用不多。

（8）组合选择题：这是多解选择题的一种变形。它把多解选择题的正确答案组合成一种固定的形式，使之成为单一性的选择题，并使难度略有提高。它的基本形式如下表：

| 供选答案编号 | A | B | C | D | E |
|---|---|---|---|---|---|
| 正确答案编号 | ①②③ | ①③ | ②④ | ④ | 其他组合 |

［范例］在四边形 ABCD 中，对角线 AC、BD 交于 O 点，AB = CD，若增加下列条件中的一个：①AO = BO；②AC = BD；③AO/OC = DO/BO；④∠OAD = ∠OBC。一定能使∠BAC = ∠CDB 成立的可选条件是　（　）

（A）②④　（B）①②　（C）③④　（D）②③④

解答组合选择题，往往采用直接法，即先不考虑备选答案，而是先从题干给的条件出发，对选项分析，然后再把分析的结果与备选答案比较，两者一致的，即为该选答案。

组合选择题综合性强，信息量大，不易猜测，其测试功能与多解选择题相同，可以测试各种层次的能力水平，也利于教师判卷。

（9）类推选择题：这类题主要检测学生是否具有一定的分析、概括问题的逻辑推理能力，要求学生根据题目所提供的条件，选出正确答案。

［范例］镭是元素周期表中第七周期ⅡA 族元素，下列关于镭的性质描述中不正确的是　（　）

（A）在化合物中呈 +2 价

（B）单质常温下跟水反应放出氢气

（C）氢氧化物呈两性

（D）碳酸盐难溶于水

（E）镭比钙金属性强

（10）信息迁移选择题：这类试题在题干中提供该学科新信息（多属最新科技成果或中学教材中未介绍过的学科知识），考生利用已掌握的学科基础知识和基本技能，经过现场阅读和一系列思维活动，进而解决题问的新问题，获得题解。这类试题主要用于

测试考生的自学能力。

[范例] 最近，科学家研制得到一种新分子，它具有空心的类似足球状的结构，分子式为 $C_{60}$。下列说法正确的是　　(　　)

(A) $C_{60}$是一种新型的化合物

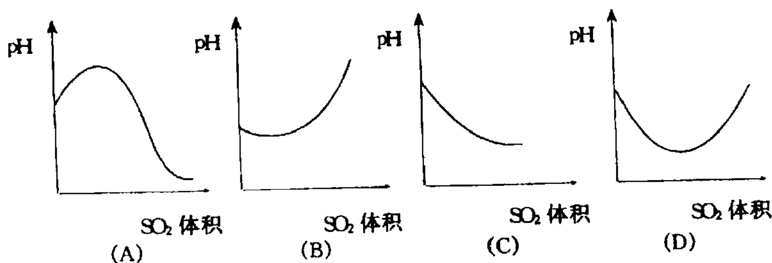
(B) $C_{60}$和石墨都是碳的同素异形体

(C) $C_{60}$中含离子键

(D) $C_{60}$的分子量是720

(11) 识图选择题：这类试题要求考生根据给定的图像来回答问题，其主要功能是考查学生的观察和分析能力。

[范例] 室温下，向饱和的 $H_2S$ 溶液中缓慢通入过量的 $SO_2$ 气体。下图所提供的是溶液的 pH 值随通入的 $SO_2$ 气体体积变化曲线示意图，合理的是　　(　　)




(12) 实验选择题：这类选择题要求考生根据题干或选项中所给仪器、药品、操作描述或示意图等选择答案，主要测量考生的实验能力。

[范例] 欲检验 $C_3H_7Cl$ 中是否含有氯元素，正确操作步骤是　　(　　)

（A）加入 $AgNO_3$ 溶液，再加稀 $HNO_3$

（B）与 NaOH 溶液共煮后，再加 $AgNO_3$ 溶液

（C）加入氯水和 $CCl_4$ 共同振荡后看 $CCl_4$ 层颜色

（D）与 KOH 溶液共煮后，用稀 $HNO_3$ 中和，再加 $AgNO_3$ 溶液

（13）计算选择题：这类试题的特点是，对一个简单的计算题列出几个计算结果，其中只有一个或两个是正确的，其余都是考生由于概念不清或考虑问题不全面常常出现的错误答案。与其他类型选择题相比，计算选择题的迷惑性选项应更具有诱惑性，否则其答案会易于猜测，使试题失去应有的价值。

[范例] 将 pH 值为 8 的 NaOH 溶液与 pH 值为 10 的 NaOH 溶液等体积混合，溶液中氢离子浓度最接近于　　　　（　　）

（A）$\frac{1}{2}(10^{-8}+10^{-10})$ mol/L

（B）$(10^{-8}+10^{-10})$ mol/L

（C）$(1\times10^{-4}-5\times10^{-5})$ mol/L

（D）$2\times10^{-10}$mol/L

（14）匹配选择题：这类选择题实际是一对一的相关选择题，供选答案可以不止一次的使用。匹配题比较适合于测试对多个知识点之间的关系的理解。其最大优点是能在比较短的时间内，测量大量相关材料，编制也比较容易。它的缺点是难以测量高层次的学习水平。

[范例] 下表左栏中只有一项和右栏中的四项相关，其余项都不与右栏中四项相关，将左栏中这一项和右栏中跟它不相关的一项选出来　　　　（　　）

| 左栏 | 右栏 |
| --- | --- |
| | ①具有臭鸡蛋气味 |
| （A）盐酸 | ②在溶液中全部电离 |
| （B）亚硫酸 | ③具有还原性 |
| （C）氢硫酸 | ④不稳定 |
| | ⑤能被浓 $H_2SO_4$ 氧化 |

（二）供答型试题

这是考生根据题干、题问自己提供题解的试题。供答型试题又可以分为两大类。

1. 限制性供答试题

考生提供题解受题干、题问强烈限制的试题，评分时很少受评分者主观因素的影响，属于客观性试题的范畴。主要包括完形填空题、释义题、简答题等。

（1）完形填空题：只有题干、题问，故意将题解突出，造成形式的不完整。强调结论，不强调解题过程。如填空题、填表题和填图题。

①填空题。它一般由删去了关键的词、字的句子所组成。它可以是一句、几句或一个段落组成。其内容一般为重要的概念、原理、时间、名人名句、实验现象、基本计算。一题可以空一处，也可以空几处。填空题具有考核面广，编题和阅卷花时短，便于考查考生对基础知识的识记程度和文字表达能力等功能。目前，填空题考查的深度和能力层次主要是了解、理解和一般性应用，对于高层次学习水平的测量以及对问题的独到见解等试题还较少见，有待于进一步研制。

②填表题。它一般由删去关键的词、字、数字后的不完整表格组成。如下表：

| 符号 | 核电荷数 | 核内质子数 | 核外电子数 |
|---|---|---|---|
| $S^{2}$ | 16 | | |
| K | | 19 | |
| $Al^{3+}$ | | | 10 |

③填图题。要求考生根据题意和图示所涉及到的有关知识，填写考核内容。

（2）释义题：提供题干和题问，需要考生提出题解。强调解题过程，同时也强调结论。如名词术语解释、释图题等。

名词解释，一般要求考生解释重要的概念、定律等。释图题是给出图式要求学生应用有关学科知识进行解释。如滴定中滴入溶液的体积与导电性大小的变化关系、溶解度大小与温度的关系等。

（3）简答题：只含题干和题问，需要考生提供简要的题解，回答题问中的“是什么”、“为什么”和“怎么样”等问题，强调解题过程，要求回答时简明扼要，抓住要点，并避免死记硬背。

限制性供答题的优点是可以直接或间接地考查学生多层次的学习水平，命题灵活，形式多样，知识覆盖面大，评分比较客观。它的不足是容易引起学生更多地注意一些学科小问题，不利于系统的、综合的掌握该学科知识。与选答题相比，答题还存在一定的书写量，不便于加大题量，同时也有一定的评分误差存在。

2. 非限制性供答试题

是在一定的题干和题问的规定下，允许考生在解题方式、内容、表达形式甚至观点上有所创造发挥的大型题。主要包括计算题、实验题、小论文题等。

（1）计算题：它的特点是从“量”的方面反映出对该学科概念的应用和对物质变化规律所掌握的程度。它不仅包括该学科知识的数学运算，并且还要通过题目所给的明显的或隐藏的已知条件找出它们与知识之间的关系。这类试题区分度大，易于把考生

不同的知识和能力的水平加以鉴别。其缺点是由于考生自由作答，评分标准不易统一。编撰时，应侧重于学科知识的综合应用，切忌故弄玄虚，偏离学科主题。有关数字应尽可能简便，数学运算不宜过难、过繁。当不需要了解计算过程时，计算题可用选择题型代替简答题型。

（2）实验题：实验题形式繁多，要求的知识面广、灵活性强、难度较大。它既能考查实验的操作与原理，又能考查物质的制取、收集、提纯、分离和鉴别，还能考查推理判断、现象描述和实验设计能力。实验题主要有：操作型、制备型、除杂型、检验型、推理判断型、定量操作型、现象描述型等。

实验题虽然在编题、阅卷方面较费时间，但由于能有效地考查学生对学科知识和实验掌握的情况，一直是理化生考卷的组成部分。近年在大型考试中，因选择题、填空题等题型日趋完善，常规实验题也有被选择填空取代之势。

（3）小论文题：这类题往往通过调查、设计、实验，综合运用“双基”解决某一学科实际问题，然后写成短文或小论文。此类试题主要考查学生综合应用知识能力、文字表达能力等高层次的学习水平。

非限制性供答题的优点在于：

第一，容易考查各种高级层次的能力。

第二，通过对解题过程的分析可以了解考生的该学科知识结构、能力分布和创造性倾向。

第三，命题花费时间少，自由度大，不仅可以从小处着手对知识进行分解式的考查，还可以从大处着眼，进行整体性的综合考查，从而达到分解式考查难以达到的目的。

非限制性供答题的缺点是：

第一，书写量大，不利于增加试题长度。

第二，评分主观性强，误差大。

## 第二节　试题的编制要领

任何试题的编制都要科学合理，试题的编制要有一定的要领，不同题型的编制要领是不同的。

### 一、二项选择题的编制要领

1. 要选择重要的概念、原理作为考点，不要单纯考核对知识的记忆，而应侧重考核对知识的理解。

2. 要保持答案的确定性和单一性，又要有一定的迷惑性。但是，题目在实质上不能模棱两可。

3. 题目中正确的个数与错误的个数应大致相近，不宜全对或全错。

4. 谬误题应尽可能避免使用“一定、全部是、总是、必然”等限定词语，这样的词意广泛，有时一看就是谬误题；正确题应少用“通常、有时、可能”等限定词，这样的论述很可能是正确的，易于猜测。

### 二、多项选择型试题编制要领

1. 尽可能采用只有一个正确答案的试题。

2. 尽可能压缩备选项的字数，而把更多的内容放在题干中。

3. 尽量使用肯定式题干，如果用否定式题干，应当用适当的形式（如加着重号）加以强调，以引起考生注意。

4. 必须使备选项构成题干的尾端，不要插在题干中间，以便于考生阅读题目。

5. 不理想选项和错误选项应具有相当的似真性和错误的典型性，以增加备选项目的迷惑程度。

6. 对题解必须避免任何内容和形式上的暗示，答案字数尽可

能相近。

7. 应该保证在任何情况下至少是三选一以上，匹配选择题要实行“差额选择”，以降低试题的猜对几率。

8. 在使用同一组选答题时，应随机改变正确答案的位置。

## 三、限制性供答试题的编制要领

1. 必须通过题干和题问对题解做出明确的限制，使答案尽可能惟一。

2. 完形题要避免那些只需机械记忆便能回答的题目，避免一题多答；当答案是数字时，要规定数量单位和数字的精确度。

3. 有可能造成多种解答的内容必须明确指出解答的方向。

4. 释义题和简答题必须对字数和答案的范围有明确具体的限制。

5. 简答题要提供可能出现的多种答案和相应的评分要求，主要用以测量较高学习水平的目标。

## 四、非限制性供答试题的编制要领

1. 在测试功能上应以考查高级层次能力为主。

2. 尽可能以能力为主线，覆盖更多的该学科知识，并注重综合性考查。

3. 保证审阅题干和题问不产生歧义。

4. 可能的情况下，将大题化为一组小题，以提高试题客观性。

5. 有必要写出详细的评分规则，尽可能减少评分误差。

6. 要给考生发挥自己真知灼见的余地。

# 第六章　考试的命题与组卷技术

命题是考试工作的中心环节。要提高考试工作的质量，关键是提高命题的科学性。命题的设计怎样才能测量出学生的知识与技能、过程与方法、情感态度与价值观等多方面的发展水平？如何将试题合理地组卷？本章将围绕这些问题进行探讨。

## 第一节　考试的命题

命题计划是试题如何编制，试卷如何组成的计划，是供命题者编制试题和试卷的依据，也可以看作是一份组卷蓝图。如果命题计划制订得准确、合理，那么，只要命题者严格按照命题计划编制试题、试卷，就可以保证试题、试卷的质量。

### 一、命题的意义

命题就是按照命题计划编制试题和试卷，为考试制造测试工具。

命题是考试工作的核心环节。要提高考试工作的质量，关键是提高命题的科学性。科学化的考试工作以命题为中心可以分为前后两大步骤：

第一步　确定考试目标——明确考试内容和方法——编制命题计划（考试蓝图）；

第二步　选择试题类型——编选试题——组卷——制订评分

标准——印卷——施考——评卷——成绩解释与使用。

第一步完成的只是考试的想法和计划，只有经过具体的命题才可能转化为可实际执行的东西；第二步是以命题的结果（试卷和评分标准）为测量工具和读数依据，来实施具体测量和处理、使用测量数据。第一步设计得再合理，计划编得再完善，如果命题工作没做好，也不会产生好的考试结果；测试过程再严密，数据处理得再科学，如果试卷编制得不好，也不可能很好地实现考试目标。所以，要做好考试工作，必须着重抓好命题环节的工作。

命题工作的重要性还表现在，许多考试的试题不仅是一定时期、一定阶段的教育和教育测量水平的反映，而且还是一种重要的历史资料，它的作用和价值并不随考试的过去而消失。做好命题工作，提高考试质量，有着比实现某次考试目标更为深广的意义。

## 二、命题的原则

命题的原则是指在考试命题过程中应当遵循的基本准则和要求。尽管考试内容不同，题型种类繁多，作用各异，但根据考试过程的基本规律，各种题型的特点，可归纳命题的一般原则：

### （一）试题抽样应具有代表性原则

具体要求是：

1. 依据课程标准和命题双向细目表命题。

2. 命题者要对总体有明确的认识，要做好考生水平的总体预测。

3. 抽取的单位应该达到足够的数量。

### （二）试题表述的简明性原则

具体要求是：

1. 命题者要具有高度的措辞技巧和准确使用词语的能力。

2. 试题的遣词、造句、行文都要经得住推敲，尽可能简明。

3. 题目格式不要使考生发生误解，要使考生明确让他做什么、怎么做、答案应以什么形式出现。

(三) 试题之间彼此独立性原则

具体要求是：

1. 每道题的含量不宜过大。

2. 各个试题不可互相牵连，不要使一个题目的回答影响另一个题目的回答。

3. 题目中不可含有暗示本题或其他题正确答案的线索。

4. 对由多人编制的试卷要逐题审订。

(四) 试题难易适度性原则

具体要求是：

1. 根据考试的性质、目的，结合考生的实际水平，确定试题难度。目标参照性考试，试题难度可低些；常模参照性考试，难度可高些。

2. 试题应有适当的梯度和恰当的区分度，试题鉴别不同水平层次考生的能力要服从考试目的。

(五) 命题与答案同步进行原则

具体要求是：

1. 命题者必须熟练掌握考试内容。

2. 明确评分标准。

3. 编制好的试题答案应妥善保管。

(六) 试题应有利于客观评分原则

具体要求是：

1. 努力提高答案的惟一性。客观性试题应有不致引起争论的答案，主观性试题对题解要有合适的范围限定和详细的评分细则。

2. 问题的正确答案，最好是有定论的，但最好不是教材上的原话。

3. 试题要以事实和原理为依据，不能脱离实际随意编选。

4. 注意各类型试题的配合使用，相互取长补短，以便能测出欲测量的知识和能力。

(七) 试题应有利于学生全面发展原则

具体要求是：

1. 树立考试既是测量工具，也是教育手段的观点。

2. 选择有利于促进学生全面发展的试题，试题内容要有一定的思想性、教育性。

3. 命题要有利于考核和促进学生提高智能。

4. 命题要注重发挥考题对学生学习方法的引导作用。

(八) 加强对命题工作科学化管理原则

具体要求是：

1. 建立健全专门的考试组织机构，加强对命题工作的指导。

2. 提高命题人员的基本素质，帮助他们掌握考试科学理论、命题原则和方法。

3. 建立考试题库。化学教师还应有计划地建立自己的小题库和试卷库。

### 三、命题计划的内容和形式

一份好的命题计划有以下两个最基本的作用：

1. 保证试题是所要测量教学内容的代表性样本，且能够反映出各个部分内容之间的相对重要性，能使试题取样适当，提高测量的效度。

2. 规划代表不同的知识内容和学习水平的各类测试项目的比例分配，在保证试题效度的同时，保证试题的难度合理。

命题计划的基本内容应根据测验的性质、目的、学生的实际水平确定，其中应明确测验的时间、题量，以及代表不同知识和不同学习水平的各类项目的比例，每种题目的分数分配与目标等。命题计划的基本形式是命题双向细目表，并可以用以下二种形式

表述：

（1）知识类别与测量水平双向细目表；

（2）试卷分数分配双向细目表；

（3）试题配方双向细目表。

下面以高一某学科第一学期期末考试命题计划为例加以说明：

**表 6－1　高一某学科第一学期期末考试命题双向细目表**

| 考试内容 | 学习水平（%） | | | | 合计（%） |
|---|---|---|---|---|---|
| | 记忆 | 理解 | 应用 | 综合 | |
| 第一章 | 6 | 8 | 11 | 5 | 30 |
| 第二章 | 2 | 12 | 12 | 6 | 32 |
| 第三章 | 8 | 14 | 12 | 4 | 38 |
| 合　计 | 16 | 34 | 35 | 15 | 100 |

根据表 6－1 规划的蓝图，再制订更加具体的试题配方表（见表 6－2）。

由表 6－2 可知试卷共由四种题型、35 个试题组成。第一种题型（可设计为填空题）共 5 个（一、1～5）；第二种题型（可设计为选择题）共 22 个（二、6～27）；第三种题型（可设计为简答或实验题）共 5 个（三、28～32）；第四种题型（可设计为计算题）共 3 个（四、33～35）。

表 6－2 也为试题的选择指明了方向。例如：项目"二、6（2）"表明为第二大题第 6 小题分值 2 分，对应的是教材第一章，属于记忆水平的教学目标，只要对照第一单元教学目标，即可找到该项目测量的具体内容。同时，此表还可说明在试题中各章知识内容以及属于各类学习水平的项目所占的比重；试题中各个知识块（章）内所涉及到的属于不同学习水平的教学的比例。

表 6－2　高一某学科第一学期期末考试试题配方双向细目表

| 考试内容 | 题号及分值 | | | | 分值合计 |
|---|---|---|---|---|---|
| | 记忆 | 理解 | 应用 | 综合 | |
| 第一章 | 一、1（2分）<br>一、2（2分）<br>一、6（2分） | 二、7（2分）<br>二、8（2分）<br>二、9（2分）<br>二、10（2分） | 二、11（2分）<br>二、12（2分）<br>二、13（2分）<br>四、34（5分） | 四、35（5分） | 30 |
| 第二章 | 二、14（2分） | 二、15（2分）<br>二、16（2分）<br>二、17（2分）<br>三、28（6分） | 二、18（2分）<br>二、19（2分）<br>二、20（2分）<br>三、29（6分） | 三、30（6分） | 32 |
| 第三章 | 一、3（2分）<br>一、4（3分）<br>一、5（3分） | 二、21（2分）<br>二、22（2分）<br>二、23（2分）<br>二、24（2分）<br>三、31（6分） | 二、25（2分）<br>二、26（2分）<br>二、27（2分）<br>三、32（6分） | 四、33（4分） | 38 |
| 分值合计 | 16 | 34 | 35 | 15 | 100 |

## 四、命题工作的任务和程序

命题的任务就是按命题计划编制试卷和试卷使用“说明书”。具体地说，命题工作有三项任务：编制至少两份具有同等效力的试卷（正卷和副本）；编写试卷的参考答案；制订试卷的评分标准。命题工作的一般程序是：

1. 学习和研究课程标准、考试大纲、命题计划，明确考试目的、性质、对象、考查哪些能力、试题的形式和数量、试题数量和分数在各部分中的分配比例、考查的重点及其他要求。

2. 按命题计划编制试题，同时写出每个试题的答案。编制试题的数量至少是需要量的二倍。写出答案的过程也就是对试题质量的检查过程，因而应与编题同时进行。

3. 对编出的试题逐一审查、修改和筛选，使通过备用的试题及其答案都科学、合理、准确。同时注明备用题的预计难度、考

查知识点及能力层次。

4. 按命题计划搭配两份以上的试卷。

5. 对各份试卷进行审查、修改和调整。

6. 制订评分标准。

试题编成后要进行试测。试测工作严格保密，最好请人试做，或由命题教师试做一遍。命题教师试做主卷的时间，以考试规定时间五分之二到二分之一为宜。

## 第二节　考试的组卷

组卷是对试题进行的优化配置，蓝图设计是组卷中首要的工作，是命题组卷的依据。蓝图设计的准确合理是做好组卷工作的前提。

### 一、什么是蓝图设计

科学的考试不仅需要制定测量目标，同时还必须进行试题和试卷的设计，而后在此基础上命题组卷，这个过程也就是组卷的蓝图设计。

### 二、组卷蓝图的形式与内容

考试蓝图一般由以下三部分组成（均为双向分布表形式）

1. 知识类别和测量水平双向细目表

2. 试卷分数分配表

3. 试题配方双向细目表

此表是在前两项工作的基础上拟定的，表的空格中具体表示了某次考核的知识点、目标和题号，有人又称之为命题计划书。见表 6－3。

表中编号 I－2（－，1）的意义为：第一知识类别（基本概念），第二知识点（物质的组成和结构），此题编入第一大题第一

小题，测量水平为“知识”。根据这张配方表，全卷共计选入45个选择题，每题2分，作为第一大题；2道题组，每题5分，作为第二大题，全卷共记100分，其中的分数应与例2的规定相吻合，此表应具有保密性。

**表6-3 高中化学毕业会考试题配方双向细目表**

| 测量目标 / 知识类别 | 知识（A） | 领会（B） | 应用（C） | 分析（D） | 综合（E） |
|---|---|---|---|---|---|
| Ⅰ基本概念 | 2题-4分<br>Ⅰ-2（-，1）<br>Ⅰ-8（-，2） | 2题-4分<br>Ⅰ-1<br>（-，21）<br>Ⅰ-6（-，4） | 1题-2分<br>Ⅰ-7<br>（-，31） | | |
| Ⅱ基本理论 | 2题-4分<br>Ⅱ-1<br>（-，13）<br>Ⅱ-7<br>（-，18） | 5题-10分<br>Ⅱ-4<br>（-，22）<br>Ⅱ-5（-，5）<br>Ⅱ-8<br>（-，23）<br>Ⅱ-12<br>（-，24）<br>Ⅱ-13<br>（-，25） | 3题-6分<br>Ⅱ-2<br>（-，3·2）<br>Ⅱ-8（-，8）<br>Ⅱ-12<br>（-，33） | 2题-4分<br>Ⅱ-1<br>（-，43）<br>Ⅱ-10<br>（-，42） | |
| Ⅲ元素化合物知识 | 2题-4分<br>Ⅲ-3<br>（-，14）<br>Ⅲ-4<br>（-，15） | 2题-4分<br>Ⅲ-2<br>（-，26）<br>Ⅲ-7<br>（-，6） | 4题-8分<br>Ⅲ-10<br>（-，19）<br>Ⅲ-11<br>（-，34）<br>Ⅲ-12<br>（-，35）<br>Ⅲ-14<br>（-，10） | 1题-2分<br>题组-5分<br>Ⅲ-10<br>（-，12）<br>Ⅲ-13<br>（=，1） | 题组<br>5分<br>Ⅲ-10<br>（=，2） |

## 三、命题中常见的问题与组卷技术

在命题工作中，经常遇到的问题是如何运用各种题型，如何确认试题的测试功能，如何控制试题的难度，如何确定考试的知识类别与测量水平的比例，如何确定试题的题量，如何将试题合理地组卷等等，这些问题的回答，实际上就是组卷技术。

### （一）题型的运用

采用什么题型以及决定它们之间的比例的基本依据是：①考试的目的和性质；②化学学科与测量内容的特点；③命题者的水

平；④考生的应试心理状况；⑤对考试后效果的考虑；⑥经济效益及其他因素。

每种题型都有其长处和局限性，因此可以将各种题型结合使用。一般而言，在大规模考试中，可以更多地采用选答型试题。这是因为：大规模考试内容涉及面广，要求有较大的覆盖面；考生人数多，希望采用机器阅卷，以提高评分的客观-致性和经济效益，如近年的中考、高考、会考，基本题型是选择题、填空题、实验题、计算题四大类，而选答型试题大约占60%。

（二）试题的考查功能

试题的考查功能是关系到考试效度的关键问题，它一般可以从两个维度进行分析，一是对考查内容的确认，二是对考查目标层次的确认。对前者的确认比较容易，而对后者的确认往往会比较困难。-般而言，考查记忆水平的试题，同教材相比，具有只是在文字、符号、结构上略有改动的特征；考查领会能力的试题，同教材相比，具有改变提问角度和需要重新排列、组合，“老问题”用新的方式来表达的特征；考查应用能力的试题同教材相比，具有在新的情景中解决问题的特征。

（三）试题难度控制

通常考试中所说的难度是指试题相对于特定的考生群体的难易程度。因此，难度控制的基本问题是对影响考生作答因素的估计水平问题。这些因素主要是①考查的知识点在化学学科中的难易程度；②教学因素，即师资水平、生源情况、教学时的重视程度、复习时的训练强化程度等；③试题本身因素，如提问的角度和措词、题干中的材料和情景对解题的提示程度、解题的思维量和思考的方向、正确答案的清晰程度或错误答案的迷惑程度等。

（四）知识内容与学习水平的比例确定

为了使试题的选择符合考试目的，要确定代表不同的知识内容、不同的学习水平或难度的各类项目的比例，使之能够对教学目标或教学内容有较好的代表性，试题的总体难度适合于被测学生。

对形成性考试而言，要以单元教学目标为依据。先要保证每个教学目标都有确定的学习水平和测试项目，而后找出那些在学习过程中容易出错的问题和应重点强化的内容，有意识地增加这类项目的比例。

对总结性考试而言，由于测试的范围更广泛一些，所以，确定各类项目的比例时更为复杂一些。代表不同知识内容的项目比例可依据该知识内容所需要的教学时间在测试范围内所有知识内容所需要的总的教学时间中的比例来确定，即达到某个教学目标所用的教学时间越多，代表该目标的测试项目所占的比例也应越大。由于教学目标多是分单元、按章节制订的，所以，可以把某章节相对应的教学时数与测试涉及所有章节教学的总时数之比，作为代表该章节教学目标的项目应占的比例。这种依据比较客观、合理，可以有效地保证测试内容取样的代表性。

确定了各类知识内容的项目比例后，还应确定各知识内容的学习水平（测量目标）。在形成性的单元测验中，各种学习水平的项目比例已经由项目代表的教学目标确定，不需要另行设计；对总结性考试而言，涉及的教学内容较多，需要对测试目标进行抽样，这就需要考虑不同难度水平的项目比例问题了。由于总结性考试要求具备对学生的学业成就进行鉴定并区别优劣的功能，因而在项目分配上，既要涉及一般水平的学习能力，还要有一定比例的较高学习水平的项目（表6－1据此设计）。对于总结性考试，不同学习水平项目比例的分配可参考表6－4。

**表6－4　总结性考试学习水平项目比例参考数值表**

| 学习水平 | 比例分配（%） | 相应通过率（%） |
|---|---|---|
| 记忆 | 15 | >85 |
| 理解 | 35 | 50～85 |
| 应用 | 35 | 15～50 |
| 综合 | 15 | <15 |

（五）试题题量的确定

试题的题量或测试内容的多少，与给定的测试时间有关。测试时间一般是在命题之前根据考试目的确定的。形成性的单元测验可用 15 分钟左右，总结性的学期或学年考试多为 90～120 分钟。测试时间要和测验目的一致，同时又要考虑到教学效率。时间过长，会占用过多的教学时间而影响教学任务的完成；时间过短，势必会限制试题的题量而影响测试目标的实现。

在测试时间确定之后，试题题量的确定依据是：学生回答试题所需要的平均时间和给定的测试时间一致。这里的平均时间是指测试者所希望的、一定比例的学生完成试题所需要的时间。一般而言，对形成性考试，要让 90% 以上的学生完成试题，以免损失过多的教学反馈信息和影响学生的学习积极性；对总结性考试，应使 80% 以上的学生完成试题，这样尽管对部分学生题量较大，但对总体学生而言，可以保证试题对测试内容有必要的覆盖面，同时提高考试的效率。

（六）试题的编排组卷

试题的编排组卷应注意以下几点：

1. 测验题的形式有多种多样，但每次测验题型不宜太多，以免造成心理干扰。为了评分客观，应尽量使用客观题型，但不排除必要时使用少量综合题型。若使用论文题型，不要过于广泛笼统，可以在总题下分几个小题来设问。

2. 题数不可太少，否则信度不高。

3. 试题的编排顺序，应根据一定的标准进行，或根据知识系统，或按照学习水平，或根据题型，无论何种都应遵循由易到难的原则。

4. 大量使用选择题的考试，应注意多类选择题的配合使用。

5. 要编写答题说明。其内容一般为：测验目的、时间限制、回答方法、记分方法等，以避免教师临场指导，尽量使测验更客

观反映考生水平。

6. 要事先编写标准答案和规定评分规则，而不是在测验后制订，否则会降低或提高标准。

评分标准一般包括三个互相联系的内容和部分。一是评分的参考答案，二是评分的办法和具体规定，三是必要的说明。如果是简答题，应按要点给分；若是论述题，可采用分解给分与整体给分相结合的办法，将大部分分数分散在各要点中，留少量分数视整体情况酌定给否。

7. 试卷的排版印刷应充分方便考生阅读作答，利于评分统分，做到紧凑美观大方。

8. 应同时准备两份以上的等值复本试卷。

## 第三节　考试题库的建立

所谓题库，就是指经过一系列的分析鉴定之后，符合一定质量标准的试题的集合。一个学科的某类考试的题库，就是能提供这个学科该类考试所用试题的试题总体。试题库的建立对于规范化的命题，科学化的考试都具有重要意义。

### 一、必要性

标准化、科学化的考试大量使用选择题，编写是很困难费时的，因此有必要预先为考试建立一个题库，在考试时随机取题组卷，以避免“闭门造题”的弊端。

### 二、建立条件

1. 所储试题的题量要足够大，试题的总体能覆盖考试大纲，要求被试者掌握学科的全部内容。

2. 试题科学合理，没有概念或逻辑的错误，并且都经过测试，

证明其质量合格，有适宜的难度和区分度。

## 三、建立方法

建立试题库按以下三个步骤实施

（一）根据知识要点和学习水平二个维度进行征题和编题。

（二）将这些征得的或由题库编制人员编制的试题在一定范围内（100－200 人次）试测，并全面回收统计。

（三）将经过试测取得质量指标合格的题按编号顺序贮存起来。

## 四、试题库示例

如表 6－5 所示。

**表 6－5　化学试题库示例**

| 试题编号 | 试题内容 | 试题答案 | 难度 | 区分度 |
| --- | --- | --- | --- | --- |
| I－5－A－1 | 摩尔是表示<br>（A）物质的量的多少<br>（B）物质的量的单位<br>（C）6.02 ×1023 个微粒<br>（D）物质质量的单位 | （B） | 0.85 | 0.30 |
| I－3－B－1 | 以下物质属于混合物的是（　　）<br>①充满氧气的玻璃管高压电击后部分转为臭氧<br>②等摩尔氯气和甲烷混合光照后<br>③漂白粉　④胆矾　⑤盐酸<br>（A）①＋②＋③<br>（B）①＋③<br>（C）②＋④<br>（D）④<br>（E）其它组合 | （E） | 0.75 | 0.31 |

这里仅举二例说明题卡的格式。其中 I－5－A－1 的意义是：

“I”是指与双向细目表相应的知识块I，即基本概念部分。“5”是指第5知识要点。“A”是指A级即“知识级的学习水平”、“1”是第1题。I－3－B－1即是第一知识块，第3知识要点，B级学习水平，第1题。

一个大型的题库应包容成千上万道测题，一道测题一经使用，五年内不应重复使用。

## 五、教师怎样建立自己的小题库

对于教师来说，建立一个自己的小题库是很有必要的。这将有助于教师科学地编制、搜集和筛选各种试题，避免教学中凭经验拼凑试卷的盲目做法，也是考试科学化的需要。

现在，教师大多都珍藏了不少试题及一些习题集子。但是，这并不等于题库。建立小题库绝不是简单的将各种试题堆积在一起，科学的方法应该是实施以下几个步骤：

### （一）制定各年级的知识要点和学习水平双向细目表

如中学化学知识一般由六类知识构成：Ⅰ、基本概念，Ⅱ、基本理论，Ⅲ、元素化合物知识，Ⅳ、有机化学知识，Ⅴ、化学计算，Ⅵ、化学实验知识。因而对各年级的知识要点和学习水平（多按美国布卢姆关于教学目标的分类方法划分），则可由有经验的教师或教研机构按课程标准及各知识要点在教材中的地位确定。

近年不少专家、学科教师在这方面的研究成果很多，教师也可以直接参照取用。

### （二）按照知识要点和学习水平两个维度编选试题，并测试其质量

教师可以将编选的试题对照上述双向细目表，对号入座，确定归属。并在每次测验后对各个题目作统计分析，以确定质量。衡量试题质量的基本指标是试题的难度和区分度。其评价标准为：

一般而言，教师可仿标准化考试的要求，将难度在0.3～0.8

之间，区分度在0.4以上的试题作为合格题选留入库。

关于试题难度、区分度的测试，教师可以利用在一个班级或一个年级（人数为100人左右最好）的各种测验后，逐题统计，求算出质量指标。

（三）将测试合格的题目制成题卡，编写贮存

在选出合格试题后，为了便于贮存、分类和使用，还应该将其制成题卡，编号入库。题卡可以制成不同式样，一般的形式为：试题全文、题型、答案及评分规定（可以写在卡片反面）、难度和区分度（如果计算不方便也可以只给出难度）、试题来源，同时将所属的知识类别、知识要点和学习水平层次与序号共同组成编号。见表6-6：

编号意义为：第六知识类别化学实验知识，第2知识点基本操作，应用级学习水平第3小题。

**表6-6 题卡示例**

| 编号 | 题目 | 答案 | 难度 | 区分度 |
|---|---|---|---|---|
| Ⅵ-2-C-03 | 用天平称量氢氧化钠时，应当把氢氧化钠放在 [ ]<br>(A) 滤纸上<br>(B) 小烧杯里<br>(C) 托盘上<br>(D) 用铝箔包好后放在托盘上 | (B) | 0.58 | 0.54 |

如果教师感到制定双向细目表的方法编号较为复杂，也可以省去第一个步骤，将题卡的编号改为考核的章节代号，如3-3-01为第三章第三节第一题。其余同上。

# 第七章　考试的组织管理

考试的组织管理即通常所说的考务工作，它的具体目标是使测试能够有步骤、有计划地进行，尽量减少和避免各种偶然因素以及工作失误对测试的干扰，保证测试结果的客观性、可靠性。同时，最大限度地降低测试过程中人力、物力和财力的消耗。健全考试制度，加强考试的组织管理，对于树立良好的考风、确保考试质量、促进考试公平都将起到重要作用。

## 第一节　考前管理工作

考前管理是考试组织管理工作的开始，搞好考前管理对确保考试组织管理质量至关重要。考前管理工作主要包括：考试工作计划的编制、命题的组织、试卷的印刷、试卷的运送和保管以及考生应考准备工作。

### 一、考试工作计划的编制

考试工作计划是主考部门对考试实施进行有效管理的手段，一份完整的考试工作计划应该包括考试目的、性质、科目及具体工作环节。为了编制考试工作计划，往往需要召开工作研究会讨论工作计划中各个项目的内容、步骤、完成时间及执行者。对于大规模的社会性考试，考试工作计划常以文件形式下达。

考试工作环环相扣，已经公布的报名时间、考试时间、成绩

公布时间及与此相衔接的取题、评卷工作时间都不能变动。因此，考试工作计划必须安排得十分谨慎、周密，某些工作安排要考虑适当留有余地。

## 二、命题的组织工作

试卷的编制即命题。命题的组织管理任务有两条，一条是为命题教师创造良好的工作条件，一条是保守试题的机密、不使其泄露。为使命题工作不受干扰，严守试题机密，最好将命题教师与外界隔离。当然这种隔离是相对的，主要是确保外界无法与命题教师取得联系。随着高科技的发展，新通讯设备在社会生活的方方面面得到广泛应用，“泄题事件”时有发生，试题的保密工作面临严峻的挑战。为此，制定命题人员守则，要求每个命题人员严格遵守，是十分必要的。

## 三、试卷的印刷工作

印刷试卷，一般交由专门的试题印刷厂承担。如果没有试题印刷厂，也要选择能与外界严格隔离的印刷厂或印刷车间来承担印刷工作。试卷后一般附答题纸，供应试者将所有考题的答案填写在上面。如果采用计算机评分，则须印制专门的答题卡，主要供应试者填写选择题的答案。另外，考区、考点、考场、考号、姓名等信息，须印在密封线以内。

印刷试卷的管理工作，主要是围绕三点来做，一是印刷和装封准确无误，二是保证在考试前印、装完毕，三是试题严格保密。

## 四、试卷的运送和保管工作

运送试题，量少的可作为绝密件通过机要机关邮寄，量多的由各考区委派专人（两人以上）取题、押运。取题时，要填写试题交接单，并由发题人、取题人分别签字。押运途中，要人不离

试题。

试题应在考试的前一天到达考点所在地，存放在安全、保密的处所，由两名以上保管员昼夜看管，直至考试。

试题在启用前如发现试卷袋缺、漏、破损等情况，立即报告主考部门，并采取有力措施，防止试题外泄、扩散。

## 五、考生应考准备工作

从考试的组织管理角度来看，考试部门在考前的考生应考准备工作主要应做好以下几个方面：

1. 发布报考简章

突然袭击式的考试，从来得不偿失。既难以考出真实成绩，又造成考生对考试的反常心理。为使考生有必要的准备时间，每次考试前都应向考生“打招呼”，说明考试的目的、内容、方法、时间及各项要求。“打招呼”的具体办法就是由各主考部门发布报考简章。

报考简章应包括如下内容：考试目的、报考对象和要求、报名地址和相关程序、考试科目和时间安排、考试方式、允许携带的物品、报考费用、查询成绩的时间和方式等及其他需要提前向考生说明的情况。

2. 颁布考试大纲

考试大纲是关于考试内容及各部分内容要求掌握程度的规定，应提早向考生公布，便于考生安排备考进度。如果考试中将采用考生不熟悉的试题类型，应在大纲中向考生详细说明此类型试题的解答办法和注意事项，并给出一定数量的样题和具体解题步骤。

实践表明，考生首次接触新型的试题，解答的速度较慢，出错的几率相对较大，心情也较紧张，考试成绩偏低。若考前能进行几次模拟考试，解题速度就会加快，并能掌握考试的规律，考试成绩明显提高。因此，考前适当进行一些熟悉题型、熟悉考规

的模拟测验，对考生来说是大有裨益的。

3. 进行考规、考纪教育

制定考规、考纪的目的是维护考场秩序，使考生能有一个良好的答卷环境，防止考生利用不正当手段获得成绩，保证测验的真实性。任何考试都应在考前向考生宣读并说明考场规则和考试纪律，重要的考试应将考规和考纪印成文字材料发给考生，并作必要的说明。以下是某省2008年普通高校招生考试考场规则[①]。

某省2008年普通高校招生考试考场规则

一、考生必须自觉服从监考员等考试工作人员管理，不得以任何理由妨碍监考员等考试工作人员履行职责，不得扰乱考场及其他考试工作地点的秩序。

二、考生凭准考证、身份证，按规定时间和准考证上各科目的考试教室、座位号参加考试。

三、考生入场，除2B铅笔、书写黑色字迹的钢笔或签字笔、直尺、圆规、三角板、橡皮外，其他任何物品不准带入考场。

严禁携带各种无线通讯工具（如寻呼机、移动电话、无线耳机）、电子存储记忆录放设备以及涂改液、修正带等物品进入考场。英语口语考试，严禁携带以上物品进入考试场所（含候考室、备考室、考场、休息室）。考场内不得自行传递工具、用品等。

四、考生应在每科开考前20分钟（第一科前移10分钟，外语科目考试在开考前35分钟）凭准考证、有效证件（身份证、现役军人身份证件）进入考场，对号入座，入座后将准考证、身份证等有效证件放在桌面靠走道边上角以便让监考员核验。考生领到答题卡、条形码和试卷后，须认真核对答题卡的张数，核对条形码上的姓名、考生号与自己准考证上的信息是否一致。如不一

① http：//www. gx－law. gov. cn/hits. asp.

致，应向监考员提出更换。在规定的时间内用黑色字迹的签字笔或钢笔准确清楚地填写答题卡上的姓名、考生号、考场号、座位号，用2B铅笔在答题卡上根据所发试卷准确填涂试题类型（A或B)，并将条形码横贴在答题卡右上角的“条形码粘贴处”栏框内。凡漏填、错填、全填或字迹不清的答卷、答题卡，导致无法正常评卷的，责任考生自负。

五、开考信号发出后才能开始答题。

六、开考15分钟后（外语科目考试开考前15分钟）禁止迟到考生进入考场，每科考试结束前30分钟起方可交卷出场（外语科目听力考试要求另行规定)，交卷出场后不得再进入考场，也不准在考场附近逗留或交谈。

七、考生在答题卡规定的区域答题。选择题用2B铅笔在选择题答题区作答，非选择题用黑色字迹钢笔或签字笔作答。选做题的信息点用2B铅笔填涂，漏涂错涂责任自负。不准用规定以外的笔和纸答题，不准在答卷、答题卡上做任何标记，否则答题卡无效。考生不准随意修改答题卡上的题号，考生必须在指定的题号里作答。凡不在指定答题题号框内作答、超出答题区域作答或擅自更改题号作答，其答案一律无效。如果解答中有画表或辅助线，先用铅笔进行画线、绘图，再用黑色字迹的签字笔或钢笔描黑确认。

八、在考场内必须保持安静、不准吸烟，不准喧哗，不准交头接耳、左顾右盼、打手势、做暗号，不准夹带、旁窥、抄袭或有意让他人抄袭，不准传抄答案或交换试卷、答题卡，不准将试卷、答题卡或草稿纸带出考场。

九、遇试卷分发错误及试题字迹模糊等问题，可举手询问；涉及试题内容的疑问，不得向监考人员询问。

十、考试终了信号发出后，立即停笔，根据监考员指令依次退出考场，不准在考场及周围逗留。

## 第二节　考场管理工作

考场管理是考试组织管理工作的核心阶段，主要包括考场编排、监考、考场秩序的维持以及巡视等环节。

### 一、考场编排

考场是考生答卷的场所，考场编排应以方便考生答卷、方便监考人员检查、方便考试秩序和考试纪律维持为原则。

每个考场以安排20到30名考生为宜。要做到单人、单桌、单行，考生前后左右之间要保持一定的距离。

来自同一单位的考生最好与外单位考生间插开编排考号。课桌的左上角（或右上角）要在考前贴上在该桌答卷的考生的考号。

每个考点都要张贴该考点考场分布的平面图，每个考场门口都要贴出在该考场分布的平面图，每个考场门口都要贴出在该考场答卷的起止考号。

### 二、监考

每个考点应设考点主任1人，副主任1至2人，负责考点的全面工作。每个考场要配备2至3名监考人员，考场外还有若干人员流动监考。

监考人员的职责是：维持考场秩序，严肃考试纪律；组织考生入场入座，分发试题和草稿纸；填写考场记录（记录缺考考生、迟到考生的考号，违纪考生的违纪情节及处理意见等）；装订答完的试卷。

监考人员的认真负责、不徇私情，是维持考场秩序和考试纪律的关键。为提高监考的责任心和业务水平，考前应对监考人员进行培训，并应制定监考守则，明确责任和工作任务。以下是某

校考试监考人员工作守则[①]。

某校监考人员工作守则

1. 每个考场必须安排两名以上监考人员。监考人员不得只由本班任课教师担任。口试必须由两位教师监考。

2. 监考人员必须提前15分钟进入考场，安排考生按指定位置就座。同时检查考生是否将工具书、参考书、笔记本和其它带字纸张带入考场，如已带入，一律放在指定地点。

3. 严格按照考试时间规定准时发卷、收卷。

4. 监考人员不得对题意和答卷要求作任何解释，若考生对试卷上的字、词、符号看不清楚，监考人员应给以说明。

5. 监考人员必须坚守岗位，认真巡视，不得擅自离开考场，不得看书看报或互相闲谈。

6. 发现违纪现象必须及时制止，并按考场规则第四条进行处理。

7. 考试结束后，要对试卷进行整理和清点，并认真填写考场记录。

8. 学校、教务处、学生处、各院（系）有关领导及职能部门工作人员要加强对考场的巡视和检查，并根据需要参加监考工作。

### 三、考场秩序的维持

考场管理，一方面要给考生正常作答创造良好的环境和条件，使考生能够发挥出平时所具有的水平，另一方面还要特别维持好考场秩序，杜绝一切可能影响考生答题的情况发生。

对于监考人员来说，维持考场秩序的工作重点主要有：根据考试科目的特点，提前准备一些答题可能用到的物品，如草稿纸、

① http：//www. xmems. org. cn/ut/bz/heh/

笔、胶布等，防止考生相互借用造成考场混乱；提醒考生听到答题信号后才能开始答题，答题时间到所有考生必须停止答题；严格监考，坚决杜绝考场舞弊现象。

### 四、巡视

巡视的主要作用是警醒监考人员、处理突发情况，学校应采取校领导、督导团、教务处多重巡视机构。巡视的频次不宜过多，巡视员负责对各考场监考人员的到位情况、履行监考职责情况、以及考生遵守考试纪律情况进行全面监督，确保各科考试的顺利进行。

## 第三节　评卷管理工作

在考试组织管理的诸多工作中，评卷管理工作是一个重要环节，也是考生最为关心的一个问题。评卷过程中要坚持客观、公正、合理的原则，最大限度地维护广大考生的切身利益。传统的评卷管理工作主要包括客观性试题的评阅和主观性试题的评阅，随着科技的发展，网上评卷也逐渐为人们所认可并接受，成为一种重要的评卷手段。

### 一、客观性试题的评阅

客观性试题或者使用计算机评分，或者采用人工评分。随着计算机的普遍使用和功能的日益完善，目前大规模考试的客观性试题评阅基本上都实现了计算机评分。答卷时随同试卷发给考生答题卡，要求考生在卡片上对他所选的答案的代号用 HB 铅笔涂抹。考后收回答案卡，输送给光电阅读机，即可评出考生的考试分数。

若采用人工评分，主考部门可组织责任心强的办事人员，对

照标准答案逐卷评阅。客观性试题评分工作简单易行，可利用各种简便办法以提高评分的速度和准确性。下边介绍的便是可行的简便办法之一：

每人只评一页试卷。评卷人按卷纸长度裁一窄条纸板，将它与卷纸对齐后，在正对各题填写答案的位置处写上标准试卷正确答案的符号。评卷时，把纸板放在卷纸上，一头对齐，于是纸板上的标准答案便与考生所选的答案逐题对应，凡与标准答案符号相同的，写“1”，与标准答案符号不同的，写“0”，逐题写好后，数出“1”的个数，填在每页的分数小计栏目内。

主观性试题的评阅较为复杂，为此有必要就评卷的组织和工作的程序加以说明。

## 二、主观性试题的评阅

1. 评卷的组织

对于大规模的考试，应在主考部门统一领导和组织下，成立试卷保管调度组、评分专家组、合分组和按题组成的各个题组(第一题组、第二题组、第三题组等)。

试卷保管调度组，负责试卷的保管、分发和收回试卷，统计工作进度。

评卷前，保管员应将每本试卷的内封皮撕下，装入封筒内，在试卷外封皮和试卷封筒相应栏目中编写上同样的号码，以便在评完后放回原封筒中。试卷代号代表的考区、考点、考场，应当保密。

调度试卷，各考区的试卷要按同一比例向各题组发放。当天发出的试卷，当天要如数收回，试卷不能在卷库之外过夜。各题组评完的试卷，要返回卷库，由调度员调往外组，各组间的试卷不能直接串换。

专家组应由命题教师和其他具有丰富教学及评卷经验的教师

组成。它的任务是：组织试评，制订评分标准的执行细则；随时抽查各题组评阅的试卷，及时发现、纠正错误及宽严不当等现象；研究解决评卷中出现的属于学科知识方面的疑难问题。

合分组负责各题都已评完的试卷的题分累计，一人合计，另一人复查，合分人、复查人都要在试卷相应栏内签名或盖章。

每个题组都只评一题，一人评分，另一人复查，评分人、复查人都要在题头的相应栏目内签名或盖章。

评卷工作应在专门的阅卷室内进行，非评卷人员不得随便进入阅卷室。

评卷中如果发现雷同试卷，应写明情况，在试卷返回卷库后，由试卷保管员查对考号，提交主考部门或主考部门委派的评卷领导小组处理。

2. 评卷的工作程序

评卷工作，通常按下述程序进行：

（1）试评

在专家组的组织下，全体评卷教师学习、讨论标准答案和评分办法，统一认识后，从各考区随机抽调部分考卷，各题组按评分标准进行试评。根据考生答卷的具体情况，专家组经与各题组研究，制订评分标准的执行细则，各题组再按细则评出若干试卷，专家组通过检查认为各题组已正确掌握了评分标准、给分宽严适当后，再转入正式评卷。

（2）阅卷给分

阅卷给分的依据是评分标准及其执行细则，原则是“客观公正，给分准确，宽严适当，前后一致”。此外，评分者要注意克服自身生理和心理因素所引起的误差，如“顺序效应”、“光环效应”、“求全效应”、“疲劳效应”等因素的影响。

（3）合分

部分试卷评完后，合分组即应开始工作，合分的主要要求是

准确无误，因此，应选调工作认真、细心、能熟练使用计算机工具的人员担当此项工作。合分后，试卷相应栏内写出的总分，必须字迹工整、规范、清楚。

（4）总复查和验收

对于重大的考试，在评卷的后期，对各题评完、合完分的试卷，还应组织复分组进行总复查。

总复查的内容，除检查有无错评、漏评、合分错误外，还应根据考试的性质确定不同的复查内容。如鉴定是否达到标准线的水平考试，对于卷面总分在标准线上下的试卷，还要检查给分宽严是否适当，因为这个分段的一两分之差，就可能关系到合格与不合格，关系到对一个人的取舍。总复查中，凡分数有变动的，总复查人都要在各变动处盖章；每本试卷复查后，都要在外封皮相应栏目内盖章。

评卷的组织有两种情况，一种是主考部门聘请教师、组织评卷，一种是委托给某所学校评卷。在后种情况下，评完卷后，主考部门应组织人员对试卷进行验收。验收的重点与总复查相同。验收合格后，试卷保管员把试卷逐本放回原封筒中，清点无误后，交给主考部门。

## 三、网上评卷

网上评卷的准确说法应该是“计算机网上辅助评卷”，这是一种新型的试卷评阅形式。网上评卷的具体操作步骤有：

1. 试卷的制作应由题签和答题卡两个部分组成。答题卡又由涂卡部分和答题部分组成。学生在答题卡上涂卡和答题，答题卡上必须粘条形码。条形码是考生的身份证，条形码要粘在指定的区域内，倒贴或是出框过多、过于倾斜计算机均无法识别。

2. 学生答题时，涂卡要用2B铅笔，笔答时要用碳素笔或签字

笔，原因是涂卡部分计算机容易识别2B铅笔的痕迹，答案扫描时用碳素笔答题清晰流畅。

3. 在互联网上申请空间，获准后将答题卡用扫描机传输到微机里，并发到互联网上指定的空间内。

4. 管理员给每位参加评卷的教师密码，评卷教师在互联网上输入密码后即可进入到所批的试题，密码具有唯一性。

5. 阅卷教师按标准答案在微机上进行阅卷。点击一次出一份试卷（只是阅卷教师所评的那道题），给出一定的分数后进行提交，一份试卷评阅结束。阅卷教师可以知道自己阅卷的数量，可以更改刚评过的一道题，余者皆不能为。

6. 所有的试卷评完后计算机提示阅卷工作结束，阅卷教师所操作的密码失去作用。计算机进行自动分数统计，统计可分为大题的分数、试卷的分数和考生所有试卷的分数。

无论是传统的阅卷方式还是计算机辅助阅卷，考生最关心的是评卷质量。实行计算机网上辅助评卷，最大的特点和优点是对每一位评卷教师的评卷质量进行实时、有效的监控。网上阅卷系统不仅设置了合理的评分误差，而且具有及时发现误差并自动提请纠正和防止误差超出设定值的功能，通过这种复评机制，将评卷员个人主观感受导致的评分误差降低到最小程度，有效防范和避免违规或马虎应付造成的误差。

## 第四节　消除考场舞弊的对策

所谓考场舞弊，就是指在考场上采取不正当的手段、想非分地得到高于自己实际水平的分数的种种违纪活动，常见的考场作弊行为有：夹带、偷看、传递、换卷、顶替以及用现代通讯工具传递答案等。

发生考场舞弊，主要原因是①：作弊考生应考动机不纯，作风不正；考试的规章制度有漏洞，使某些人有机可乘；监考失职，考场纪律不严。及时采用有效措施，消除可能导致考场舞弊的各种因素，是促进考试公平的有力保证。

## 一、堵塞漏洞，使作弊者无机可乘

制度上的漏洞是出现舞弊的客观原因，建立严密的报名考试制度，是消除舞弊的必要保证。对于社会性考试来说，以下规定和措施是不可少的：

1. 两个“三对照”。顶替和换卷，是考场上最严重的作弊，必须严防。报名时，考生至少要交两张照片，一张贴在准考证上，一张贴在按考场装订的底册上，考生入考场后，将两个照片与考生本人三者相对照，这是防止准考证换贴照片、冒名顶替的重要办法；考生答卷中，把准考证上的考号、课桌上的考号与试卷上填写的考号三者相对照，是防止填写他人考号、换卷的重要办法。

2. 混编考场，交换监考。将同单位的考生与其他单位的考生间插开，混编考场，是防止互换试卷、传递、交头接耳的必要措施。有的考试，熟人监考有一些困难，这时，应将不同地区、不同单位的监考相交换。

3. 考生单人、单座、单行，考桌间拉开一定距离，客观性试题采用“A、B 卷”。采用主观性试题，如果不是监考玩忽职守，只要考生座位间拉开一定距离，偷看题是难以奏效的。客观性试题，考生只写符号，且写在试卷的固定位置处，试卷上的答案，邻座考生是一目了然的。为此，应采用“A、B”卷。奇、偶考号的考生使用不同试卷，并使每个考生周围四个考生的考号奇偶与其相反，即考号呈如下排列：

---

① 杨德平．杜绝考试作弊的对策分析［J］．当代教育科学，2007，(2)：62.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|
| 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 |

4. 按考号顺序逐桌收试卷，装订前再按考号顺序逐页检查，这是防止、发现换写考号的重要办法。

5. 考试后试卷保存一段时间，如有人反映顶替、换卷、照抄等问题，可核对笔体和试卷解答的内容。

## 二、加大力度，增强监考人员责任感

监考人员是考试的“执法官”，是维护考场纪律的关键。在有上述的制度规定和措施的情况下，只要监考人员认真负责，不徇私情，杜绝考场舞弊是完全可能的。因此，选聘严守纪律、敢于执行纪律、铁面无私的监考人员，做好监考工作，是杜绝考场舞弊的最重要的办法。

## 三、加强巡视，向各考点派巡视员

对于重要的考试，如中考、高考、英语四六级考试、计算机考试、公务员考试等，主考部门向各考点派出一定数量的巡视员，监督、检查各项规定的执行、各项措施的落实情况，也是十分必要的。

## 四、提高警惕，严防高科技作弊①

随着现代通信技术的发展，考试作弊手段开始与高科技挂钩。高科技作弊造成了恶劣的社会影响，挑战着考试的公平公正性。

---

① 史履海．发挥监测技术手段，查获考场作弊团伙［J］．中国无线电，2008，(4)：27.

近几年，高科技作弊事件时有发生，如《焦点访谈》就曾曝光寿光考场作弊案、广东电白的 BP 机舞弊案。还有一些考生通过手机、纽扣摄像机、隐形耳机等，企图避开监考人员的视线达到作弊的目的。由于这些通讯设备具有体积小、不易被发现的特点，给监考人员的监考工作带来很大的困难。

预防和解决高科技作弊最重要最有效的措施就是严把“监考关”，高科技作弊现象无论是群体或个体行为，都与考试工作人员和监考教师有直接或间接的关系。其一，对于考场内高科技作弊的个体行为，大多数情况监考人员都能发现，“管与不管”只是主观问题。其二，高科技作弊的群体行为，能否做成的关键是试题是否泄密，而考试工作人员是渠道之一。其三，监考人员应了解当前的主流作弊手段，通过考场巡查，注意考生脸部、体态、空间信息，“看”神态变化，判断考生心理，杜绝考生通过无线电向外传送信息或提前离开考场造成试题的泄密。

## 第五节　综合实践活动课考试实施探讨

综合实践活动是新一轮基础教育课程体系中从小学三年级开始设置的必修课程，它是指以学生的兴趣和直接经验为基础，以与学生学习生活和社会生活密切相关的各类现实性、综合性、实践性问题为内容，以教师指导、协助，学生自主、合作、探究学习为主要特征的一类新型课程，在新课程实施中具有独特功能和价值①。如何组织实施综合实践活动的学业评价，是一个新的课题。

---

① 张永国．综合实践活动学生学期学业评价初探［J］．上海教育科研，2004，(9)：67.

## 一、评价实施的原则

鉴于综合实践活动课的基本特征和教学特点，对学生的学业评价应重参与、重过程，强调评价主体的多元化、评价内容的综合性和全面性、评价标准的合理性，以及评价方法、手段的多样化。因此，综合实践活动评价的原则有四条：一是参与原则，即综合实践活动注重学生亲身参与和学生全员参与，注重学生对学校提供的自主和自由的学习时间和空间的利用，强调教学计划规定的课时活动量的参与情况和参与态度，同时重视学生自我评价并自觉参与评价。二是过程原则，即综合实践活动评价特别关注学生参与活动的过程，重视对过程的评价和在过程中的评价，并且把对学生的评价与对学生的指导紧密结合起来。三是综合原则，综合实践活动本身由“信息技术教育、研究性学习、社区服务与社会实践以及劳动技术教育”四部分组成，各部分的教学目标又包含各自的认知、情感、技能和能力等要求。因此，综合实践活动评价的内容、要求、过程和结论都应是综合而全面的。四是发展原则，鼓励学生发挥自己的个性特长，施展自己的才能，激励学生积极进取、勤于实践、勇于创新，不断促进学生综合素质的发展。

## 二、评价的内容与形式

综合实践活动评价的内容通常应涉及以下几个方面：参与活动的课时量与态度，在活动中所获得的体验情况，知识、方法、技能的掌握情况，创新精神和实践能力的发展情况，活动的收获与成果[1]。

评价的形式多样，一般可采取“四结合”的形式，即教师的

---

1　孙岩．走进综合实践活动课程［J］．中小学教学研究，2004，（3）：8.

评价与学生的自评、互评相结合，对小组的评价与对组内个人的评价相结合，对书面材料或考试的评价与学生的报告、活动、展示的评价相结合，定性评价与定量评价相结合、以定性评价为主等形式。

## 三、评价的主要方法①

总体上说，综合实践活动课评价在模式上侧重于过程评估、质性评定、学生表现评定。在收集评价资料的途径上主要有观察、答辩、交流、录像、记录案例、学生的自我反思等。

评价的方法多种多样，要使综合实践活动课评价的方法科学合理，恰当地选择评价方法是非常重要的。在综合实践活动课评价过程中，评价的方法要多样化，评价的方式要灵活。

### （一）教育鉴赏

美国教育评价专家艾斯纳（E. W. Eisner）在20世纪80年代就提出把评价对象当艺术品一样，运用自己的生活体验和直接智慧，去欣赏、体验、感受、把握评价对象的动态变化特征。它强调教育评价要向生活世界回归；强调多元价值取向，允许学生选择自己所喜欢的一种表现形式来展现自己的所学；强调评价应揭示学生是如何解决问题的；强调对活动动态结果的把握。在综合实践活动课评价时采用教育鉴赏与教育评价的方法时必须注意：一是评价要由单一的教师评价学生，转变成学生自己评价自己、学生评价学生和学生评价老师的多向评价，营造一种和谐安全的环境，切实促进学生的发展；二是要倡导积极的评价，教师和学生要学会恰当的赞赏别人，赞赏每一位学生的独特性、兴趣爱好和专长，赞赏每一位学生所取得的哪怕是极其微小的进步。评价者要把评价定位在学生成功的方面，多进行赞赏评价。

---

① 梁燕玲．综合实践活动课评价的理论思考［D］．广西师范大学，1999：22－29.

（二）学生成长记录与分析

综合实践活动课的评价是一项长期的任务，教师不能就事论事评价某一活动中学生的表现，而应该将学生看成一个不断发展的个体，要将评价贯穿于学生的整个发展过程。因此，学生成长记录与分析在综合实践活动评价中也有一定的应用。

1. 记录与分析的内容

在对学生成长记录与分析时，首先要对成长记录内容进行选择。选择成长记录内容时，要建立明确的评价目标，与综合实践活动课程紧密结合，并能反映学生的成长变化。其次，学生的成长记录和分析内容要涵盖：学生原有的知识、技能、能力；学生原有的学习经历、生活经历；学生的需求、想法；学生侧面的发展情况；学生需要被关注和帮助的方面；有关帮助学生的措施、方法等。

2. 评价主体与评价手段

学生成长记录与分析的主体是多元的，包括教师、学生、家长以及其他学生的评价。而评价的方式也是多种多样的，例如观察、访谈、行为检查表、作品与日常表现分析、活动表现分析等。

（三）研讨评价法

研讨评价法是学生表现评定的有效方法之一，是质性评价的典范，也是综合实践活动课评价的一个不可缺少的方法。采用这一评价方法时需要注意以下几点：

1. 研究问题的质量

一个高质量研究问题的提出是非常重要的，它将直接影响到研究的质量和进展。好问题往往不期待单一或最适合的答案与预期反应，而是学生感兴趣的、开放式的、探索式的，能够不断引发对话和学生对学习内容更深入的理解。因此，在采用研讨评价法对学生综合实践活动进行评价时，确定高质量的研究选题是评价实施的重要环节。

2. 选择研讨评定的要点

研讨评定的要点主要包括：关注学生提出问题、解决问题的能力；关注学生的创新能力；考察学生独立完成任务的能力；关注学生与他人的合作、交流与分享；考察学生是否具备了基本的科学研究的能力；评价学生综合运用知识的能力。

## 四、评价实施应注意的问题

一是要坚持多元化的评价取向。综合实践活动课非常强调将学生视为真正的学习主体，认为学生是学习过程的思考者、活动者、体验者、合作者和建构者。同时，非常注重学习过程中主体价值的多元化与个性化，尊重每一位学生的个性特点。因此，在综合实践活动课评价的方式上也必须超越传统的忽视中间过程，注重目标取向的结果性评价，致力于构建一种将教学过程与教学情景融为一体的多元化评价[①]。

二是倡导质性评价与量化评价的整合。评价标准的量化即分数，是传统评价的主要手段。分数评价学生的学业有一定的参考价值，但不具有肯定的价值。由于学生在综合实践活动课堂中，所涉及的不仅仅是最后学到了多少知识，更重要的是在学习过程中情感的投入、态度的转变、方法的习得和生活的体验，这些变化需要在评价中得到充分的体现和评定。因此，将质性评价与量化评价有机地整合，才能达到全面评价学生的目标。

三是采取以小组成绩为基础，兼顾个人的评价形式。综合实践活动课的内容是广域性的，主要以课题为载体，学生也通常以小组合作的形式开展学习。在每个小组中，学生要从事各种需要合作与帮助的活动。这种学习方式更能突出小组中每位学生的主体地位，培养主体参与意识，激发学生的创造潜能。因此，综合

---

① 刘兰英．小学探究性学习评价的思考［J］．教育发展研究，2001，（10）：34－35.

实践活动课的评价应考虑采用对合作小组的评价与对组内个人的评价相结合，强调对小组的评价。在对小组评价的基础上，根据个人的表现，对小组成员进行等级评价。评价可分为四个等级，第一个等级是在活动中能出色完成任务起到骨干作用的成员，或者是有创新表现的同学；第二等级是实践能力较强，能够圆满完成任务的成员；第三等级是能够完成自己分工任务的成员；第四等级是参加活动不积极或不能完成课题组交给任务的成员。评价者可根据具体情况对小组成员进行评定。

# 第八章　考试的质量统计与分析

一个科学的考试，离不开对考试结果采用科学统计方法的分析，也就是采用教育统计学、教育测量学及计算机科学的理论和方法去对考试分数进行整理、转换与解释。本章联系中学化学学业测验的实例，对科学的统计与分析化学成绩的方法作一简介。

## 第一节　考试分数的初步整理——图表法

考试后获得的分数往往是不系统的、零乱的，很难以从这些分数中找出规律去解释分数。因此，首先必须对考试分数作初步的科学的分组、归纳、概述，使之系统化，初步整理数据的基本方法是统计表法和统计图法，即图表法。

### 一、统计表

统计表是系统地陈述和表现统计资料的表格。它具有简明清晰、条理清楚、易于比较、便于查对等优点，是表达数字资料的一种重要方式。

（一）制表的基本原则和要求

1. 制表的基本原则

（1）表的结构要简单明了，一张表只能有一个中心，说明的问题要十分明确，避免绘制臃肿的、包罗万象的大表；

（2）表的层次要清楚，项目的排列要按照逻辑顺序合理安排。

2. 制表的要求

（1）标题：标题即表的名称，应确切、简明、扼要地说明表的内容，标题应写在表的上方，必要时，应在标题下注明资料的来源（地点、单位和时间）。

（2）标目：标目即表格的项目

标目按在表内的位置可分为横标目和纵标目，横标目所指明的数字在同一横行，位于表的左侧，纵标目所指明的数字在同一纵行，位于表的右侧，必要的时候，可在横标目和纵标目的上方加上适当的总标目。

标目的内容又可分为主语和谓语，主语是统计表叙述的内容，谓语是用以叙述的统计指标；习惯上，一般地主语放在横标目位置上，谓语放在纵标目的位置上，若主语较复杂时，可移动一部分到纵标目上去；若谓语较复杂时，可移动一部分到横标目上去。一个设计良好的统计表，循“主语－谓语－数字”自左向右的顺序阅读，可以形成一个通顺句子。

（3）线条：线条不宜过多。顶线、底线、隔开纵标目与数字的横线是表的三条基本线条，如设有总和，可在总和的上方添一横线，其余线条尽量减少。特别是表的左上角不宜有斜线。

（4）数字：表内数字必须准确，且一律用阿拉伯字母表示，位次对齐，小数的位数一致。表内不应有空格，暂缺或未记录可用“…”或“－”表示，数字若是“0”，实际为“0”或数值极小则应填写“0”。

**表 8－1　统计表的基本格式**

△△△△　（标题）　顶线↖

| 横标目的总标目（亦可空白） | 纵标目（谓语） |
| --- | --- |
| 横标目（主语） | 数字（用阿拉伯字母） |

底线↗

具体统计表一例：

表 8－2　某班某学科成绩统计表

| 组别 | 优 | 良 | 中 | 差 |
|---|---|---|---|---|
| 一 | 4 | 7 | 7 | 2 |
| 二 | 2 | 5 | 8 | 3 |
| 三 | 1 | 9 | 9 | 3 |

表 8－2 可以按“主语－谓语－数字”，自左向右的顺序阅读：“某班某学科成绩的统计结果，第一组同学优等的有 4 人，良等的有 7 人，中等的有 7 人，差等的有 2 人；第二组……”

下面我们用××中学××学年度高一化学两个班 103 人的上学期期末考试成绩为例，具体说明统计表的制法。下文中不另说明的也均以此数据为例。

表 8－3　××中学××学年度高一两个班 103 人上学期化学期末考试成绩表

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 73 | 83 | 84 | 84 | 86 | 69 | 85 | 64 | 91 | 84 |
| 78 | 87 | 63 | 83 | 57 | 62 | 60 | 67 | 73 | 85 |
| 91 | 81 | 72 | 79 | 75 | 74 | 90 | 89 | 97 | 54 |
| 83 | 77 | 80 | 78 | 72 | 74 | 90 | 79 | 63 | 78 |
| 71 | 66 | 74 | 80 | 77 | 70 | 65 | 79 | 88 | 45 |
| 49 | 77 | 72 | 93 | 79 | 79 | 57 | 74 | 85 | 72 |
| 80 | 84 | 70 | 80 | 47 | 79 | 80 | 81 | 61 | 84 |
| 83 | 77 | 98 | 87 | 91 | 61 | 92 | 84 | 80 | 76 |
| 84 | 80 | 74 | 78 | 92 | 73 | 74 | 70 | 85 | 78 |
| 79 | 63 | 46 | 64 | 90 | 87 | 63 | 76 | 58 | 88 |
| 68 | 85 | 44 | | | | | | | |

（二）频数分布表的列法

频数分布是指一群分数中每一个分数值（或一组数据）所出现次数的分布情况，一般将这群分数按一定标志和一定顺序排列，即成为频数分布。当数据较多时，将全部数据按大小分成若干小组，把所有的数据归到有关的组内，得到数据在组内出现的次数

（亦称频数），从而清楚地看出全部数据的分布情况，这种分组往往制成表格的形式，即为频数分布表。

从考试或测验中所获得的原始分数，如表8－3，它们是杂乱无章的，一经编制成频数分布表后就一目了然了。

编制频数分布表的方法与步骤如下：

1. 求全距（R）

找出原始分数中的最大值和最小值，全距＝最大值－最小值，如表8－3数据的全距。

$$R = 98 - 44 = 54$$

2. 决定组数和组距

组数是指将全距分成组的个数，组距是指每一组内包含的间距（用i表示）。组数一般以10～20组为宜，分组过多会达不到化繁为简的目的，分组过少则易产生信息失真。

$$\text{组距}(i) = \frac{\text{全距}(R)}{\text{组数}}$$

表8－3数据以11组为宜　$i = \frac{54}{11} = 5.4 \approx 5$，组距一般取整数值。

在学校成绩统计中，一般按组距为5或10进行分组即可。

3. 决定组限

组限即每组的起止范围。每组的最低值为下限，最高值为上限，由于数值较大一组的下限就是数值较小一组的上限，因此可省略各组的上限，如“44～49”简写成“44～”。

4. 求组中值（Xc）

组中值又称组中点，是居于该组分数尺度上正中心位置的数值。其求法：

$$\text{组中值 } X_c = \frac{\text{下限} + \text{上限}}{2}$$

或

$$\text{组中值}(X_c) = \text{下限} + \frac{\text{组距}}{2}$$

后一公式在运用时较为方便，特别是组距不是定值时，如 44 ~49 这一组的组中值：

$$X_c = \frac{44+49}{2} = 46.5$$

或

$$X_c = 44 + \frac{5}{2} = 46.5$$

5. 登记频数（f）

分好了组后就可以将数据按所属的组一个一个地登记于表中，登记的符号可自选。登记完毕，算出各组频数（f）列入频数列。在登记频数时，不要重复登记组限值，且标准要一致。

6. 登记累积频数

即从数值最小的一组开始，每降下一组，必须把以上各组的频数累积起来，登记进去，列入累积频数列，含最大值一组的累积频数应等于总频数。

7. 转换累积频数为频数百分比（P）

即将各组的累积频数除以总频数再乘以 100，在比较几组分数而各组的总频数又不相等时，就要将各组的频数转化成百分比。

说明：在列频数分布表时，累积频数和累积频数百分比可视具体需要而列，不必千篇一律。

表 8 -4 就是运用频数分布表整理表 8 -3 的化学成绩。

**表 8 -4　103 人高一化学成绩频数分布表**

| 组别 | 组中值（$X_c$） | 画记 | 频数（f） | 累积频数 | 频数百分比（p） |
|---|---|---|---|---|---|
| 44 - | 46.5 | 止 | 4 | 4 | 3.88 |
| 49 - | 51.5 | 一 | 1 | 5 | 4.85 |
| 54 - | 56.5 | 止 | 4 | 9 | 8.74 |
| 59 - | 61.5 | 正下 | 8 | 17 | 16.5 |
| 64 - | 66.5 | 正一 | 6 | 23 | 22.3 |
| 69 - | 71.5 | 正正丅 | 12 | 35 | 34.0 |

| 组别 | 组中值（$X_c$） | 画记 | 频数（f） | 累积频数 | 频数百分比（p） |
|---|---|---|---|---|---|
| 74 - | 76.5 | 正正正下 | 18 | 53 | 51.4 |
| 79 - | 81.5 | 正正正正 | 20 | 73 | 70.9 |
| 84 - | 86.5 | 正正正下 | 18 | 91 | 88.3 |
| 89 - | 91.5 | 正正 | 10 | 101 | 98.1 |
| 94 - | 96.5 | 丅 | 2 | 103 | 100 |

（资料来源：表8－3）

## 二、统计图

统计图是以几何图形的形式表达统计数量关系的重要工具，它能生动、形象、直观地反映统计数量的分布情况、差异细节及变动趋势。为便于分析起见，一般在运用统计图表达数据关系时，附有统计表。

常用的频数分布图有直方图和频数多边图，下面分别介绍它们的画法。

### （一）直方图

直方图是用面积表示频数分布，用各组上下限矩形面积表示各组频数，其画法如下（以表8－4为例）：

1. 画横轴

把表8－4中上下限或各组中值分别置于横轴上，作图时，注意在横轴的两端至少各空出一个组距的位置。

2. 画纵轴

纵轴表示频数，注意要标明尺度及其单位。

3. 画直条

按各组频数定出在纵轴上的高度，在各组频数高度处画一横线，横线宽度等于组距宽度，再在各组内通过上、下限各画两条纵线与标志各组频高度的横线交成一个矩形，即成直方图，如图8－1。

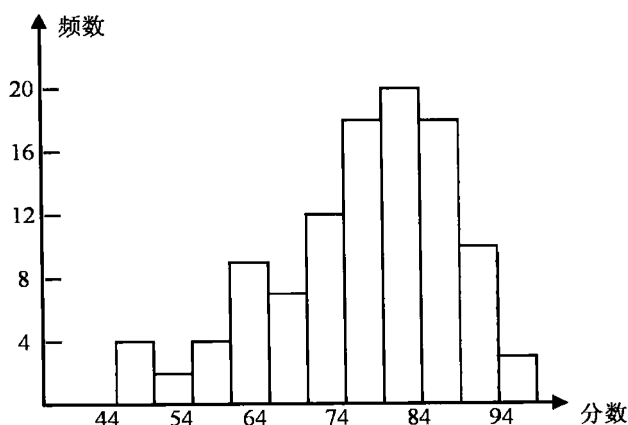



**图 8－1　103 名高一学生化学期末考试成绩直方图**
（资料来源：表 8－4）

注意：由于横轴组距变化是连续的，故各矩形间不能留有空隙，但可以省去每个矩形的内侧垂线。如图 8－2。

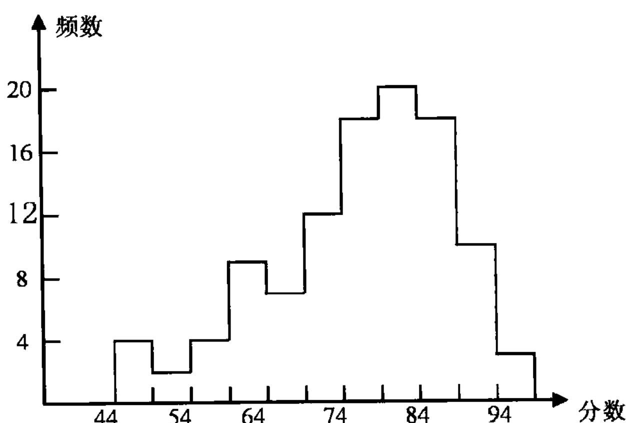



**图 8－2　103 名高一学生化学成绩直方图**
（资料来源：表 8－4）

4. 为了直方图的美观，图的纵轴高度与横轴宽度之比为 3∶5

为宜。

（二）频数多边图

频数多边图的画法与直方图画法相似，在频数多边图中，以每组的组中值代表该组的数据，作为横坐标，再在纵坐标上找出相应的频数，相交成一点，然后把每个点用直线连续起来，即成频数多边图，图形的两端应外推至外侧一组的组中点与基线相连（即频数为0）。

据表8－4数据绘制频数多边图，如图8－3。

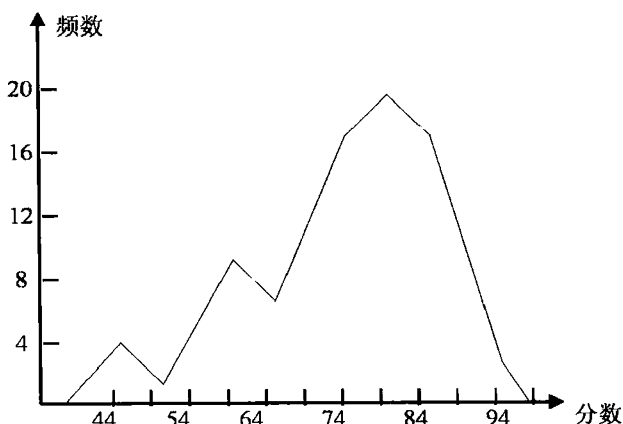



**图8－3 103名高一学生化学成绩频数多边图**
（资料来源：表8－4）

从上图可以看出，频数多边图能够较清晰直观地反映统计数据的分布情况及变化趋势，与直方图比较，它能给人以清晰地轮廓。但频数多边图不及直方图精密，因为直方图能确切地代表每一组内的频数。在比较两组或两组以上数据的分布情况时频数多边图较直方图清楚，因为直方图内线条较多，容易发生重叠而模糊。基于这一点，频数多边图得以广泛的应用。在进行这种比较时，将不同的频数分布图绘制在同一图上。此时应注意，两者组

距必须相同，如果在一张图上绘制两个总频数不同的频数分布多边图时，须将各组的频数转化成百分比，然后以频数的百分比为高度制图。

下面请根据表8-5、表8-6、表8-7练习绘制频数分布多边图。

**表8-5　××中学××学年度高一（1）班上学期化学期中考试成绩**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 73 | 66 | 61 | 91 | 59 | 56 | 56 | 47 | 75 | 65 |
| 76 | 69 | 68 | 61 | 46 | 63 | 80 | 46 | 61 | 64 |
| 47 | 91 | 84 | 85 | 46 | 86 | 79 | 71 | 68 | 74 |
| 82 | 64 | 63 | 88 | 71 | 58 | 71 | 74 | 78 | 62 |
| 66 | 41 | 56 | 75 | 76 | 41 | 67 | 41 | 68 | 43 |
| 82 | 23 | | | | | | | | |

**表8-6　××中学××学年度高一（6）班上学期化学期中考试成绩**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 67 | 78 | 62 | 72 | 74 | 60 | 68 | 60 | 69 | 71 |
| 63 | 80 | 79 | 76 | 44 | 60 | 36 | 71 | 47 | 75 |
| 72 | 54 | 58 | 68 | 58 | 46 | 71 | 81 | 91 | 46 |
| 70 | 41 | 52 | 66 | 55 | 66 | 70 | 48 | 58 | 63 |
| 40 | 68 | 58 | 81 | 65 | 39 | 60 | 67 | 65 | 62 |
| 30 | | | | | | | | | |

**表8-7　××中学高一（1）、（6）两班化学成绩频数分布表**

| 组别 | 组中值（Xc） | 画记 | | 频数（f） | |
|---|---|---|---|---|---|
| | | 高一（1） | 高一（6） | 高一（1） | 高一（6） |
| 20- | 22.5 | 一 | | 1 | 0 |
| 25- | 27.5 | | | 0 | 0 |
| 30- | 32.5 | | 一 | 0 | 1 |
| 35- | 37.5 | | 丅 | 0 | 2 |
| 40- | 42.5 | 𠄌 | 下 | 4 | 3 |
| 45- | 47.5 | 正 | 𠄌 | 5 | 4 |
| 50- | 52.5 | | 丅 | 0 | 2 |
| 55- | 57.5 | 正 | 正 | 5 | 5 |
| 60- | 62.5 | 正丅 | 正下 | 7 | 8 |

| 组别 | 组中值（Xc） | 画记 | | 频数（f） | |
|---|---|---|---|---|---|
| | | 高一（1） | 高一（6） | 高一（1） | 高一（6） |
| 67 - | 67.5 | 正正 | 正正 | 10 | 10 |
| 70 - | 72.5 | 正 | 正丅 | 5 | 7 |
| 75 - | 77.5 | 正一 | 正 | 6 | 5 |
| 80 - | 82.5 | 下 | 下 | 4 | 3 |
| 85 - | 87.5 | 下 | | 3 | 0 |
| 90 - | 92.5 | 下 | · | 2 | 1 |

（资料来源表 8-5，8-6）

（三）常见的化学成绩频数分布曲线类型

当所观察的成绩分数个数较少时，绘成的频数分布图表现出不规则的齿形分布，如图 8-3 的频数多边图。当我们所观察的成绩数目足够多时，则齿形分布逐渐变成光滑而有规则的曲线。

1. 单峰对称型

这种类型的分布特点是“中间多，两边少”，根据峰的高度与宽度又可分为正态分布、高狭峰和低阔峰三种。正态分布是在数理统计中广泛应用的连续型概率分布，它符合一定的函数表达式，高狭峰是相对于正态分布显得高而狭的峰，低阔峰是相对于正态分布显得低而阔的峰，其图形关系如下：

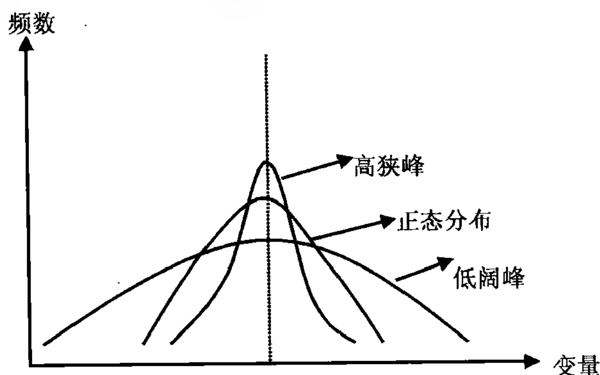



图 8-4　单峰对称型频数分布曲线

2. 单峰非对称型

这就是通常所说的偏态分布，它可分为正偏态和负偏态，与正态分布曲线相比，频数分布偏向左的为正偏态，频数分布偏向右的为负偏态，如图 8－5 所示。

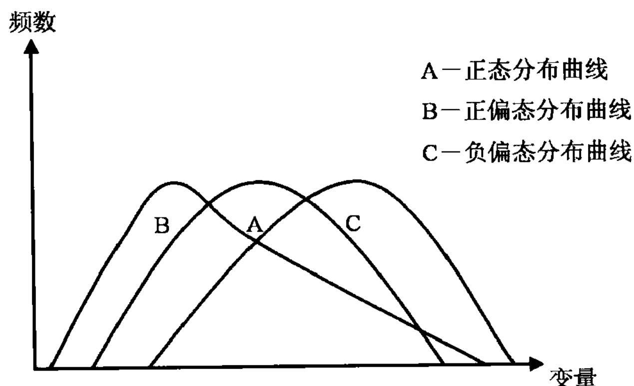



图 8－5 单峰非对称型频数分布曲线

3. 多峰分布型

多峰分布型频数分布曲线即频数分布图上出现多个峰值（图 8－6）。有些成绩分布曲线由于组距过小而会出现多峰分布，若加大组距则多峰消失，呈单峰分布。

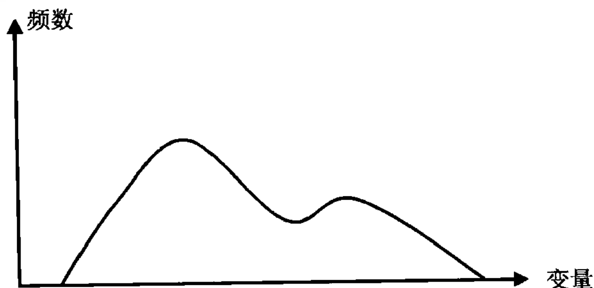



图 8－6 多峰分布型频数分布曲线

一次考试成绩的结果，考生成绩的分布以哪种分布最为理想？这要根据所进行考试的目的而定。

（1）常模参照性考试：如高中入学考试、高考，考生的水平应该是正态分布。如果呈现严重的偏态分布，则可能有二种原因：一种是考生的水平是正态分布，但考试题目过分容易或难，评分过松或过严，造成考试不能如实地反映考生的水平。考试过难或评分过严，致使中等以上考生得了低分，呈现正偏态分布；反之，考试过易或评分过松，呈现负偏态分布。另一种原因是考生的实际水平分布确实是呈偏态分布，但这种情况极少。

又如，进行国际奥林匹克竞赛、全国竞赛等各种知识竞赛，主要是选拔出成绩特别好的学生，且名额有限，因此希望统计成绩呈现正偏态分布。

（2）目标参照性考试：如教师自编试题的期中考试、期末考试，单元测验和章节测验等，要求学生尽可能多地掌握教学目标所规定的内容，因而希望学生成绩呈负偏态分布，越偏向负偏态分布，说明学生掌握的知识程度越高。

总之，在考察某项考试的成绩分布情况时，不能一概而论，应视具体情况进行具体分析，从而获得自己所需的信息。

## 第二节　考试分数的进一步整理——特征量数法

经过归组、列表和绘图等初步整理工作，对考试分数的分布情况有了一个总体的印象，但是，若要对分数进一步的分析和评价，必须将其以量的形式表达出来，才能体现出其科学性。因此，经过初步整理的数据必须进行进一步的数量化的描述。在教育统计中，总是用这些分数的集中趋势和离散程度来描述。

在教育统计学中，集中量是代表一组数据的典型水平或集中趋势的量，它能反映频数分布中大数据向某一点集中的情况。差

异量是表示一组数据变异程度或离散程度的量。在描述一组数据的频数分布时，必须同时用集中量和差异量来综合分析，才能获得较完全的信息，下面分别介绍集中量和差异量的有关知识。

## 一、集中量

常用的集中量有算术平均数、中位数、众数、加权平均数等，现分述如下：

### （一）算术平均数（$\overline{X}$）

算术平均数是所观察值的总和除以总频数所得的商，简称平均数，用 $\overline{X}$ 表示。它是统计学中最常用的一种集中量的指标。其计算方法有以下三种。

1. 原始数据计算法

计算公式：$\overline{X}=\frac{\sum X_i}{N}$

式中　$X_i$——每个考生的分数

$\sum X_i$——全体考生分数总和。

N——考生的总人数

如某小组 12 人期中考试某学科成绩分别为：

$$\overline{X}=\frac{67+78+62+72+74+60+68+60+69+71+63+80}{12}=68.7$$

同样可计算表 8－3 分数的算术平均值 $\overline{X}=75.9$，这里通常运用计算器将数据一一输入，再按算术平均数 $\overline{X}$ 键也可显出。

此公式一般用于分数较少的情况。

2. 组中值计算法

当学生较多时，通常先将原始分数（即考试直接获得的分数）编制成频数分布表，然后把频数分布表中的组中值近似地看成每组分数的算术平均值。那么每一组的分数之和就可以用该组的组中值乘以该组的频数而求得，而所有分数之和就可用各组分数之和来代替，然后再除以总频数，即得算术平均值。

计算公式：

$$\overline{X}=\frac{f_1X_1+f_2X_2+\cdots+f_kX_k}{f_1+f_2+\cdots+f_k}=\frac{1}{N}\sum_{i=1}^{K}f_iX_i$$

$f_i$——表示第 i 组分数的频数；

$X_i$——表示第 i 组组中值；

$\sum f_iX_i$——表示各组组中值与该组频数乘积之和；

N——表示总频数。

用组中值法计算平均数的步骤如下表：

**表 8－8　××中学高一化学成绩算术平均数组中值法计算表**

| 组别 | 组中值 $X_i$ | 频数 $f_i$ | 组中值与频数之积 $f_iX_i$ | 利用公式 $\overline{X}=\frac{\sum f_iX_i}{N}$ 计算算术平均值 |
|---|---|---|---|---|
| 44－ | 46.5 | 4 | 186 | |
| 49－ | 51.5 | 1 | 51.5 | |
| 54－ | 56.5 | 4 | 226 | |
| 59－ | 61.5 | 8 | 492 | $\sum f_iX_i=7884.5$ |
| 64－ | 66.5 | 6 | 399 | $\sum f=103$ |
| 69－ | 71.5 | 12 | 858 | $\overline{X}=\frac{7884.5}{103}=76.5$ |
| 74－ | 76.5 | 18 | 1377 | |
| 79－ | 81.5 | 20 | 1630 | |
| 84－ | 86.5 | 18 | 1557 | |
| 89－ | 91.5 | 10 | 915 | |
| 94－ | 96.5 | 2 | 193 | |

（资料来源：表 8－4）

直接由表 8－3 的原始分数计算出平均值为 75.9，相差 0.6 分，相对误差为 0.8%，可见用频数分布表计算的平均数与其由真值计算结果非常接近，误差较小。

3. 简化值计算法（或代号计算法）

这是在组中值法基础上更为简化的方法，它是运用算术平均值的以下二点性质。

（1）若每一个原始分数（$X_i$）加上（或减去）一个常数（A），那么算术平均值 $\overline{X}$ 的变化也同样加上（或减去）这个常数，即：

$$\frac{\sum (X_i \pm A)}{N} = \overline{X} \pm A \qquad (\overline{X} = \frac{\sum X_i}{N})$$

（2）若每个原始分数（$X_i$）同时乘以（或除以）一个常数（A），那么算术平均值的变化也同样是乘以（或除以）这个常数，即：

$$\frac{\sum AX_i}{N} = A\overline{X} \text{或} \frac{\sum X_i/A}{N} = \frac{\overline{X}}{A} \qquad (\overline{X} = \frac{\sum X_i}{N})$$

简化值计算法的基本原理是把各组的组中值（X），同减去一个常数 A（这个常数称为假定均数）然后同除以另一个常数 i，（这个常数就是组距，但必须在组距相等的情况下，才能够这样做）。其结果就是组中值的简化值（d）。

$$d = \frac{X - A}{i}$$

由于简化值计算的平均值显然被缩小了，据上面算术平均值的性质可知，真实算术平均值应为：

$$\overline{X} = A + \frac{\sum fd}{\sum f} i$$

式中 d——各组组中值的简化值；

f——各组的频数；

i——组距。

（3）现仍用表 8－4 数据为例来计算平均值。

1）列计算表，将表 8－8 中的组别、组中值和频数分别转入表 8－9：

2）选择假定均数（A），一般用频数较多组且位于中间的组中值为假定均数。本例 A＝76.5。

计算简化值（d）：　　$d = \frac{X - A}{i}$

简化值很有规律，在 d＝0 上下，组中值小于假定均数的各组 d 值依次为－1、－2、－3，……，组中值大于假定均数的各组 d 值依次为 1、2、3……故 d 可直接写出，不计算。

表 8－9 103 名高一化学成绩简化值法计算平均值

| 组别 | 组中值 X | 频数 f | 简化值 $d=\frac{X-A}{i}$ | fd | 计算 |
|---|---|---|---|---|---|
| 44－ | 46.5 | 4 | －6 | －24 | |
| 49－ | 51.5 | 1 | －5 | －5 | $\sum fd=1$ |
| 54－ | 56.5 | 4 | －4 | －16 | $\sum f=103$ |
| 59－ | 61.5 | 8 | －3 | －24 | $A=76.5$ |
| 64－ | 66.5 | 6 | －2 | －12 | $i=5$ |
| 69－ | 71.5 | 12 | －1 | －12 | |
| 74－ | 76.5 | 18 | 0 | 0 | $\overline{X}=76.5+\frac{1}{103}\times5$ |
| 79－ | 81.5 | 20 | 1 | ＋20 | $=76.5$ |
| 84－ | 86.5 | 18 | 2 | ＋36 | |
| 89－ | 91.5 | 10 | 3 | ＋30 | |
| 94－ | 96.5 | 2 | 4 | ＋8 | |

（资料来源：表 8－4）

3）求简化值的算术平均数，即$\frac{\sum fd}{N}$。本例：

$$\frac{\sum fd}{N}=\frac{1}{103}$$

计算原分数的算术均数

$$\overline{X}=A+\frac{\sum fd}{N}\times i=76.5+\frac{1}{103}\times5=76.5$$

由此可见，用简化值计算的结果与用组中值法计算结果完全相等，而前者计算简便得多。

（二）中位数（Md）

中位数是依一定顺序排列的一组分数中央位置的数值，在这一数值上、下各有一半频数分布着，可见中位数是将一组数据一分为二的数值，它也是一种表示集中量的指标，但它划分的界限不一定与平均数划分的界限相吻合。

中位数的计算方法如下：

1. 原始数据计算法

将一组分数依大小顺序排列后，如总频数为奇数，就以位于

中央的分数作为中位数；如总频数为偶数，则以最中间的两个数据的平均值为中位数。如以下 12 个数 60、60、62、63、67、68、69、71、72、74、78、80 的中位数 $Md=\frac{68+69}{2}=68.5$（$\overline{X}=68.7$），而 30、40、63、67、70、72、78 这 7 个数的中位数 $Md=67$（$\overline{X}=60$）。

可见，中位数不一定同该组分数的算术平均数相等，中位数的这种求法只宜用于少量数据。

2. 频数分布表计算法

当数据较多时，可先编制频数分布表，再按下列步骤求中位数（以表 8－4 数据为例）

（1）求 N/2，本例 N＝103，N/2＝103/2＝52。

（2）确定中位数所在组，由上往下（或由下往上），累积频数，直至略大于 N/2 为止，该组就是中位数所在的组，表 8－4 中 74～79 这 ·组就是中位数所在的组。

（3）确定由中位数所在的组取多少个频数就能使由上往下或由下往上的累积频数等于 N/2，即求 $N/2-n_1$；（$n_1$ 为小于中位数所在组的下限的频数和），表 8－10 中 $N/2-n_1=52-35=17$。

（4）计算从中位数所在组所取频数的距离，即求 $N/2-n_1)\frac{i}{f_{md}}$（$f_{md}$是中位数所在组的频数），本例（$N/2-n_1$）$\frac{i}{f_{md}}$＝（$103/2-35$）$\times\frac{5}{18}=4.6$。

（5）求中位数：$Md=L+\left(\frac{N}{2}-n_1\right)\times\frac{i}{f_{md}}$（L 为中位数所在组的组下限），本例 $Md=74+4.6=78.6$。

从上面计算可知，103 名学生分数高于 78.6 的有 51 名（或 52 名），低于 78.6 的有 52 名（或 51 名），见表 4－10。

若从下往上计算则用下列公式：

$$Md = u - (N/2 - n_2)\ i/f_{md}$$

式中 u——中位数所在组上限；

$n_2$——大于中位数所在组上限频数总和。

当由上往下计算中位数时，如果小于某组下限的累积频数正好等于总频数的一半，那么该组的下限就是中位数；当从下往上计算时，大于某组上限的累积频数正好等于总频数的一半，那么该组的上限就是中位数。

**表 8－10　103 人高一化学成绩中位数表**

| 组别 | 频数 f | 累积频数 | 计算中位数 |
|---|---|---|---|
| 44－ | 4 ↑ | 4 | |
| 49－ | 1 | 5 | |
| 54－ | 4 | 9 | $L = 75$ |
| 59－ | 8 $n_1 = 35$ | 17 | |
| 64－ | 6 | 23 | $Md = L + (N/2 - n_1) \times \frac{i}{f_{md}}$ |
| 69－ | 12 ↓ | 35 | |
| 74－ | 18 | 53 | $= 74 + (\frac{103}{2} - 35)\ \frac{5}{18}$ |
| 79－ | 20 ↑ | 73 | |
| 84－ | 18 $n_2 = 50$ | 91 | $= 78.6$ |
| 89－ | 10 | 101 | |
| 94－ | 2 ↓ | 103 | |

（资料来源：表 8－4）

（三）众数（Mo）

众数就是一组数据中频数出现最多的那个数值。它也是集中量的一个指标，一般情况下可用观察法求得，或用公式求得。

1. 用观察法直接寻找粗略的众数

（1）在一组原始数据中频数出现最多的那个数据就是众数，这种方法同样只宜用于少量数据。

（2）在频数分布表中，频数最多一组中值就是粗略众数，当两个相邻的组频数都相等，且最多时，则这两组的分组点就是众

数，如表 8-4 的频数最多的一组是 79-84，其组中值为 81.5，故该表数据的众数 Mo=81.5。

2. 用公式求理论众数的近似值

从理论上说，众数是与频数分布理论曲线最高点相对应的横坐标上的一点，但众数求法甚繁，一般用英国的统计学家皮尔逊（K. Pearson）发现的算术平均数 $\overline{X}$、中位数 Md 与众数 Mo 三者之间的关系的经验公式来求理论众数的近似值。

（1）在原始分数分布呈现正态分布或接近正态分布时，有公式：

$$Mo \approx 3Md - 2\overline{X}$$

由前面推算，对于表 8-4 数据，Md=78.6　$\overline{X}=76.5$，则 $Mo \approx 3Md - 2\overline{X} = 3 \times 78.6 - 2 \times 76.5 = 82.8$，这种方法与观察法获得的粗略众数（81.5）相近。

（2）在整个频数呈现偏态分布时，应采用以下公式来进行比率调整：

$$Mo = L + \frac{f_a}{f_a + f_b} i$$

i——组距；

L——众数所在组的组下限；

$f_a$——大于众数所在组上限那个相邻组的频数；

$f_b$——小于众数所在组下限那个相邻组的频数。

下面是某次某学科竞赛成绩，由于成绩分布呈偏态，故必须进行比率调整，见表 8-11。

（四）加权平均数（$\overline{X}_t$）

加权平均数是几个作用比重不同的算术平均数（或个别代表值）的平均数。它是在特定的条件下使用的平均数。

计算公式：

$$\overline{X}_t = \frac{\sum N_i \overline{X}_i}{\sum N_i}$$

$N_i$——表示第 i 组的频数，又称权数；

$X_i$——表示第 i 组平均值。

当数据彼此之间在总平均数中的比重不同时，可用加权平均数来计算。

表 8－11　某次某学科竞赛成绩众数表

| 分数 | 频数 | 求众数 $Mo = L + \frac{f_a}{f_a + f_b} i$ |
|---|---|---|
| 42－ | 4 | |
| 45－ | 9 | |
| 48－ | 20 | $L = 48 \quad f_a = 14$ |
| 51－ | 14 | $f_b = 9 \quad i = 3$ |
| 54－ | 7 | |
| 57－ | 5 | $Mo = 48 + \frac{14}{14+9} \times 3 = 49.83$ |
| 60－ | 4 | |
| 63－ | 2 | |
| 66－ | 1 | |

如某校规定，学科成绩的计分比例为平时：期中：期末＝3：3：4，某学生的平时成绩为 87 分，期中考试成绩为 69 分，期末考试成绩为 73 分，则该学生总平均数为：

$$\overline{X}_t = \frac{87 \times 3 + 69 \times 3 + 73 \times 4}{3 + 3 + 4} = 76 \text{（分）}$$

当已知各班的平均分数，求算全年级的总平均分时，不能把各班平均分加起来除以总班数，这是因为各班的人数不等，计算总平均分要用加权平均数，尤其在两班人数相差很大的情况下，用加权平均数与用算术平均数求得的结果相差很大。

如甲、乙两班学生，甲班 42 名化学成绩平均分为 88.9，乙班 57 名成绩平均分为 74.3。求甲、乙两班化学的总平均分数。

用加权平均数计算：

$$\overline{X}_t = \frac{42 \times 88.9 + 57 \times 74.3}{42 + 57} = 80.5$$

用算术平均数计算：

$$\overline{X}=\frac{88.9+74.3}{2}=81.6$$

可见二者平均分相差竟达 1.1 分，误差较大。

算术平均数、中位数、众数和加权平均数都表示一组数据集中的趋势，这些集中量是被研究数据的代表数值，具有典型性。集中量是成绩在数量分布上的最重要的标志，其中加权平均数的用途广泛，作用突出，是最好的集中量。

为便于记忆与比较，现把这四个集中量列表如表 8－12。

**表 8－12　四个集中量的比较**

| 名称 | 符号 | 计算公式 | 优点 | 缺点 | 适用范围 |
|---|---|---|---|---|---|
| 算术平均数与加权平均数 | $\overline{X}_t$<br>$\overline{X}_t$ | $\overline{X}=\frac{\sum X_i}{N}$<br>$\overline{X}=\frac{1}{N}\sum f_i X_i$<br>$X=A+\frac{\sum fd}{N}i$<br>$X_t=\frac{\sum N_i\overline{X}_i}{\sum N_i}$ | 严密可靠，代表性强，易计算，易理解 | 易受极端值影响，使其失去代表性 | 使用最广，利用频数分布表计算时，组距必须是确定的 |
| 中位数 | Md | $Md=L+\left(\frac{N}{2}-n_1\right)\frac{i}{f_{md}}$<br>（从上→下）<br>$Md=u-\left(\frac{N}{2}-n_2\right)\frac{i}{f_{md}}$<br>（从下→上） | 意义明确不受两极端数据影响 | 缺乏灵敏性，不如平均数可靠，不能用代数方法计算 | 量数的频数很多，频数又有较集中的趋势。 |
| 众数 | Mo | $Mo\approx 3Md-2\overline{X}$<br>$Mo=L+\frac{f_a}{f_a+f_b}i$ | 不受极端值影响 | 不可靠，不能用代数法计算 | 适用粗略估计分组数据末端组限不清时（无法计算）可用众数代表集中趋势，在统计学中用作集中量的价值很小 |

（五）用集中量描述频数分布曲线的对称性。

（1）当一组分数的平均数、中位数和众数相等时，即 $\overline{X}=Md=Mo$，成绩呈正态分布。

（2）当 $\overline{X}-Mo>0$ 时，成绩呈正偏态分布。

（3）当 $\overline{X}-Mo<0$ 时，成绩呈负偏态分布。

分布曲线的不对称性可用偏斜度（或偏态量）来描述。

公式：$$S_k=\frac{3\ (\overline{X}-Md)}{S}\quad (S-\text{标准差})$$

当 $S_k=0$ 时成绩呈正态分布。

$S_k>0$ 时成绩呈正偏态分布。

$S_k<0$ 时成绩呈负偏态分布。

一切资料都是有些偏斜的，但除非明显得很，一般都把它当作正态分布，例如若 $\overline{X}=39.80$　$Md=40.42$　$S=23.2$，

$$\text{则}S_k=\frac{3\ (39.80-40.42)}{23.2}=-0.08$$

由于在这个频数分布中 $S_k=-0.08$ 几乎接近于 0，因而可以说是接近正态分布的。

## 二、差异量

让我们先来分析下列二组化学成绩。

甲组：72、72、66、60、54、78、48

乙组：90、84、72、70、60、48、36、30

$\overline{X}_{甲}=60$　　　　$\overline{X}_{乙}=60$

可见两组的平均分相等，能不能说这两个组的化学成绩完全一样呢？仔细观察上述数据可以看出，甲组学习成绩最大值与最小值相差不多（72－48＝24），乙组相差很大（90－30＝60），这说明乙组学习程度参差不齐，即两组平均分相同，但离散程度不同。可见，仅有 $\overline{X}$ 还不够完整描述两组分数，还需要一个用来表示离散程度的参数，即差异量来描述。

下面就介绍常用的几种差异量指标：全距、四分位距与标准差。

(一) 全距 (R)

全距是一组数据中最大值与最小值之差，又称极差。全距值越大，表示数据的个别差异或离散程度越大，如上例甲组全距 R = 72 - 48 = 24，乙组 R = 90 - 30 = 60，可见乙组的个别差异比甲组大。归组资料的全距一般是最大一组与最小一组的组中值之差，如表 4 - 4 的全距 R = 96.5 - 46.5 = 50。

全距的意义明确，计算简单，但由于它没有考虑到中间数据的差异，因而代表性不强，易受极端值影响，因而它通常用作辅助参数指标。

(二) 四分位距 ($Q_D$)

四分位距是第 3 四分位数与第 1 四分位数差的一半。四分位数是指把一组分数从小到大排列或编成频数分布表后，把总频数分成相等的四段，那么第 1 四分位数相当于量表上第 25% 的那一点，第 3 四分位数相当于量表上第 75% 的那一点，四分位距 ($Q_D$) 的计算公式：

$$Q_D = \frac{Q_3 - Q_1}{2}$$

式中 $Q_1$—第 1 四分位距；

$Q_3$—第 3 四分位距。

其计算方法如下：

1. 原始数据求算法

如有 16 个数据 25、22、29、12、40、15、14、39、37、31、33、19、17、20、35、30，其四分位距计算如下：

先将这 16 个数从小到大排列后求 Q1 和 Q3。

12、14、15、17、19、20、22、25、29、30、31、33、35、

Q1 = 18　　　　Md = 27

37、39、40



Q3 = 34

$$Q_D = \frac{34 - 18}{2} = 8$$

2. 频数分布表计算法

先用下列公式求出 Q1、Q3：

$$Q_1 = L_{Q_1} + \left(\frac{N}{4} - n_1\right) \frac{i}{f_{Q_1}}$$

$$Q_3 = L_{Q_3} + \left(\frac{3}{4}N - n_2\right) \frac{i}{f_{Q_3}}$$

$f_{Q_1}$、$f_{Q_3}$分别表示 $Q_1$、$Q_3$ 所在组频数，$L_{Q1}$、$L_{Q3}$分别表示 $Q_1$、$Q_3$ 所在组下限，$n_1$、$n_2$ 分别表示 $Q_1$、$Q_3$ 所在组下限频数总和。

对表 8－4：

$$Q_1 = 64 + \left(\frac{103}{4} - 23\right) \times \frac{5}{6} = 66.3$$

$$Q_3 = 84 + \left(\frac{3}{4} \times 103 - 73\right) \times \frac{5}{18} = 85.2$$

$$Q_D = \frac{85.2 - 66.3}{2} = 9.45$$

四分位距一般用于中位数表示集中量时来表示分数的差异程度，由于它不受两极端分数的影响，因而较全距可靠，但运算不方便。

（三）标准差（S 或 σ 表示）

标准差是最重要、最常用的差异量指标，在计算过程中避免两极端数值的影响大大超过全距，在统计学中得到广泛应用，它常用于算术平均数为集中量的场合，相互配合。

1. 标准差的定义公式：

$$S = \sqrt{\frac{\sum (X_i - \overline{X})^2}{N}}$$

式中　N——总频数；

$\overline{X}$——算术平均数；

$X_i$——第 i 个数据。

标准差可以直接地、概括地、平均地描述变异的大小，对于同质资料来说，标准差越小，表明数据的变异程度越小，即数据越整齐，分布范围越集中；标准差越大，表明数据的变异程度越大，即数据越参差不齐，分布范围广。

2. 标准差的计算方法

（1）原始数据计算法

在数据不多的情况下，把上面的定义式进行变换成如下形式，用于计算即方便又精确：

$$S = \sqrt{\frac{\sum X_i^2}{N} - \left(\frac{\sum X_i}{N}\right)^2}$$

如下面 5 个数据的标准差为：83、87、86、81、88

$$S = \left[\frac{83^2 + 87^2 + 86^2 + 81^2 + 88^2}{5} - \left(\frac{83 + 87 + 86 + 81 + 88}{5}\right)^2\right]^{1/2}$$

$$= 2.61$$

这里，亦可用通常的计算器先把原始数据输入，然后按标准差键即可显示。

（2）简化值计算法

在数据较多时，可先制频数分布表，再用简化值计算标准差：

$$S = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} \cdot 1 \qquad \left(d = \frac{X - A}{i}\right)$$

其表格与表 8－9 相比只多 $fd^2$ 一列。

3. 标准差的组合

在学校、教育统计行政部门统计工作中，常常统计各校（或班级）的平均分和标准差，或者统计量繁多，不可能将全部数据统一计算，只能先分成几部分，然后再汇总等等，这些就涉及到

标准差的组合问题，即如何由各组成部分的标准差求算总体标准差的计算问题，这个问题可用下列公式解决：

$$S_t=\sqrt{\frac{f_1\ (S_1{}^2+X_1{}^2)\ +f_2\ (S_2{}^2+X_2{}^2)\ +\cdots+f_n\ (S_n{}^2+X_n{}^2)}{f_1+f_2+\cdots+f_n}}$$

$$=\sqrt{\frac{\sum f_i\ (S_i{}^2+X_i{}^2)}{\sum f_i}}$$

式中：

$S_t$——组合标准差；

$S_i$——各组标准差；

$X_i$——各组平均数与总平均数之差；

$f_i$——各组的频数。

例如：某校高一四个班语文成绩如表 8－14，若求全年级学生语文成绩的总标准差，其方法可见表 8－15。

**表 8－13　103 人高一化学成绩标准差计算表**

| 组距 | 组中值（X） | 简化值 $d=\frac{X-4}{i}$ | 频数 f | fd | $fd^2$ | 计算标准差 |
|---|---|---|---|---|---|---|
| 44－ | 46.5 | －6 | 4 | －24 | 144 | |
| 49－ | 51.5 | －5 | 1 | －5 | 25 | |
| 54－ | 56.5 | －4 | 4 | －16 | 64 | $A=76.5$ |
| 59－ | 61.5 | －3 | 8 | －24 | 72 | $\sum fd^2=555$ |
| 64－ | 66.5 | －2 | 6 | －12 | 24 | $\sum fd=1$ |
| 69－ | 71.5 | 1 | 12 | －12 | 12 | |
| 74－ | 76.5 | 0 | 18 | 0 | 0 | $S=\sqrt{\frac{555}{103}-(\frac{1}{103})^2}\times 5$ |
| 79－ | 81.5 | 1 | 20 | 20 | 20 | $=11.6$ |
| 84－ | 86.5 | 2 | 18 | 36 | 72 | |
| 89－ | 91.5 | 3 | 10 | 30 | 90 | |
| 94－ | 96.5 | 4 | 2 | 8 | 32 | |

（资料来源：表 8－9）

**表 8 – 14　高一四个班语文成绩表**

| 班级 | 人数（f） | 平均分（$\bar{X}$） | 标准差（S） |
|---|---|---|---|
| 1 | 45 | 75 | 9 |
| 2 | 38 | 78 | 8 |
| 3 | 40 | 69 | 10 |
| 4 | 37 | 82 | 7 |

**表 8 – 15　四个班语文成绩组合标准差 St 的计算表**

| 组别 | f | $\bar{X}_n$ | $S_n$ | $S_n^2$ | $X_n$ | $X_n^2$ | 计算 |
|---|---|---|---|---|---|---|---|
| 1 | 45 | 75 | 9 | 81 | -0.8 | 0.64 | $\bar{X}=\frac{f_1\bar{X}_1+f_2\bar{X}_2+\cdots+f_n\bar{X}_n}{f_1+f_2+\cdots+f_n}=75.8$ |
| 2 | 38 | 78 | 8 | 64 | 22 | 4.84 | $S_t=\sqrt{\frac{f_1(S_1^2+X_1^2)+f_2(S_2^2+X_2^2)+\cdots+f_n(S_n^2+X_n^2)}{f_1+f_2+\cdots+f_n}}$ |
| 3 | 40 | 69 | 10 | 100 | -6.08 | 46.24 | $=\sqrt{\frac{45(81+0.64)+38(64+4.84)+40(100+46.24)+37(49+38.44)}{45+38+40+37}}$ |
| 4 | 37 | 82 | 7 | 49 | 6.2 | 38.44 | $=9.8$ |

注意：在求算全年级化学成绩标准差时，不能将各班标准差相加再除以班数，那样，计算的结果是各班标准差的平均数，而不是全年级化学成绩的总标准差。

（四）差异系数（CV）

标准差的重要功能之一是用来比较不同频数分布之间的离散程度，但是，若遇到两种数据的测量单位不同，如两门不同学科的考试，或两种数据的单位相同但两者的平均数相差很大，如同一次化学测验，那么这种数据之间的差异程度就不能用标准差来比较，必须用差异系数来表示。

全距、四分位距及标准差等都是带有与原始数据相同单位的名称，我们称之为绝对差异量，而差异系数是相对差异量，它是标准差与算术平均值的百分比，这是一个没有单位的相对量，计

算公式为：

$$CV = \frac{S}{\overline{X}} \times 100\%$$

差异系数大，表示数据的离散程度大；差异系数小，表示离散程度小。

注意：用差异系数表示的离散程度的大小，并不能表明哪个好，哪个差，要判断优劣还要结合具体问题的具体要求而定，简单地运用差异系数的大小来说明事物的优劣是不妥当的。

差异系数的主要功能和用途有：

1. 比较不同单位资料的差异程度

如同次不同科考试分数的比较，见表 8－16。

**表 8－16　某校高一年级期中考试成绩**

| 科目 | 平均分（$\overline{X}$） | 标准差（S） | 差异系数（CV） |
|---|---|---|---|
| 化学 | 76.5 | 7.8 | $CV = \frac{7.8}{76.5} \times 100\% = 10.20\%$ |
| 物理 | 82.5 | 6.8 | $CV = \frac{6.8}{82.5} \times 100\% = 8.28\%$ |

可见化学成绩比物理成绩差异性要大些，物理成绩稍显整齐。

2. 比较单位相同而平均分数相差较大的两组数据的差异程度

如同科不同班考试分数，见表 8－17。

**表 8－17　高一两班某学科成绩**

| | 平均分（$\overline{X}$） | 标准差（S） | 差异系数（CV） |
|---|---|---|---|
| 甲班 | 89 | 9.24 | 10.38% |
| 乙班 | 67 | 8.02 | 11.96% |

可见，这两班的成绩差异程度不大，但乙班不如甲班成绩整齐。

3. 比较班内学习分化的程度——制作成绩分数比较图

如高一（2）班7次化学测验成绩，见表8－18。

**表8－18　××××高一（2）班7次化学测验成绩**

| 统计量 | 测验的次数和内容 | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | 氧化还原 | 卤族元素 | 摩尔 | 期中考试 | 硫、硫酸 | 期末考试 | 碱金属 |
| X | 87.14 | 80.35 | 82.69 | 71.97 | 68.85 | 72.84 | 77.80 |
| S | 14.59 | 8.89 | 13.00 | 12.03 | 12.97 | 9.89 | 10.56 |
| CV（%） | 16.74 | 11.06 | 15.72 | 16.72 | 18.89 | 13.58 | 13.57 |

统计资料表明：CV值一般分布在5～35%之间，如果CV值大于35%时，可以怀疑所求平均数是否失去意义；如果CV值小于5%时，可以怀疑平均数计算是否有误。在班内或年级中，某科（次）成绩CV≤9%，可以认为无分化现象；若CV≥20%，则说明分化现象严重；而当9%＜CV＜20%时，可以认为该科（次）学习上有分化的苗头，应当重视。

图8－7为成绩分化比较图。

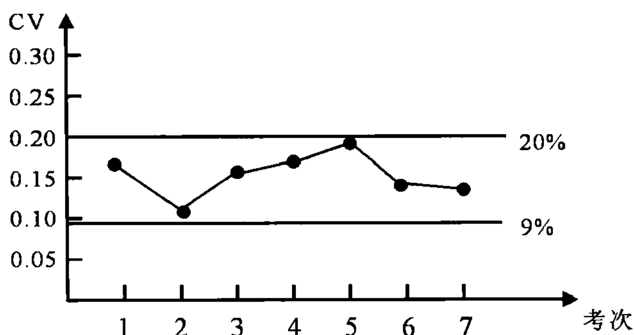



**图8－7　成绩分化比较**

由上图可见，7次化学考试成绩分布均在9%＜CV＜20%之

间，表明班内化学学习没有较大的分化，但已有分化的苗头，其中第三章《硫、硫酸》学习成绩分化比其它几次更明显，这些信息应引起化学教师注意，找出原因和相应的对策，以利于大面积提高化学教学质量。上图还可以用于某次考试各科成绩的比较，了解各科成绩的差异。

通过对考试分数的进一步描述，我们从量的方面对分数进行了考察，数据的集中趋势最常用的是平均数，它最能代表某组数据的总体水平。同时，我们又常用标准差来反映数据的差异程度。在考察一组数据时，必须同时考察其集中趋势的差异程度，两者相辅相成，才能进行进一步的分析。

小结：平均数和标准差是学校统计中常用的两个统计量，有了它们基本就能反映全部考试分数的情况，尤其是在人多的场合，又准确又方便，不必再划分分数段了，这一部分内容要重点掌握平均数和标准差的意义与计算。

利用平均分和标准差分析学生成绩，一般有四种情况：

平均分相同或接近，标准差各异；

平均分相同或接近，标准差也相同（或接近）；

平均分不同，标准差相同；

平均分不同，标准差也不同；

无论哪种情况，标准差只表明数据的分散情况，不能用标准差的大小判断成绩的优劣。

平均分、标准差都可利用电子计算器的统计档 SD 进行计算。

## 第三节　原始分数的转换与解释

原始分数（Raw Scores）是考试直接得到的分数。学校常用的传统记分法（分等法、百分制记分法）都是以原始分数直接解释成绩，这种方法简便易行（主要以平均分为标准），但缺乏科学

性，这与原始分数的特点有关。

## 一、为什么要对原始分数进行转换

### （一）原始分数的特点

第一，原始分数的单位没有普遍意义，而且不等距。所谓单位的普遍意义，就是相同单位应该在人们心目中有相等的价值。如1m，不管在什么人心目中都一样长。可是相同的分数的价值却因评分者不同、学校不同而不同。所谓等距，是指单位相等。如每一厘米长度相等，100cm与90cm之差，一定等于10cm和20cm之差。分数是以分为单位，但每一分却不一定相等。例如某题难度较大时，全对为10分；某题难度较小时，全对也是10分，这两个10分的价值显然不等。

第二，原始分数参照点不同。所谓参照点，就是测量事物时计算的起点。参照点即使不同，但只要意义明确，大家公认，也可以作为测量时计算的起点。但测验或考试评分的参照点却有很大分歧，评分者之间很少一致。由于各自评分的参照点不同，所以原始分数之间无法进行比较，而不比较就不能确定学生成绩的优劣和在集体中的相对地位了。

因此，原始分数的最大不足就是在于缺乏可比性，具体又表现在以下三点：

（1）学生某次考试所得原始分数与其它次考分不具有可比性；

（2）不同学科的考试成绩不具有可比性；

（3）不同学科的考试成绩不具有加和性。

### （二）原始分数转换的作用

原始分数因单位不等距和没有以绝对零点为参照点，所以意义不明确，不能用代数方法处理，不同测验的原始分数之间也不能进行相互比较，所以必须将其转换。

原始分数的转换，就是把原始分数转换成具有一定参照点和

单位的量表分数，这种分数也称导出分数（Derived Scores），其作用在于：

（1）使每个人的分数与同一测验的所有分数的分布联系起来，以明确每个分数的位置，从而获得每个分数在团体中的相对地位的信息。所以，这种导出分数过去也称地位量数。

（2）使不同测验的分数有了比较的可能性。

（3）有的导出分数能进行“+、-、×、÷”符号运算，能用代数方法处理，为进一步统计分析打下基础。

导出分数主要有两类：一类是在分布中保持分数间距离的比例关系的导出分数；另一类是不保持这种比例关系的导出分数。Z分数、T分数等由原始分数通过线性转换而得到的导出分数属于前一类；百分等级分数和标准九分数等属于后一类。我们主要介绍如何将原始分数转换成标准分数Z分数和T分数。

## 二、怎样把原始分数转换成标准分数

标准分数（Standard Scores）有Z标准分和T标准分等。

（一）Z标准分（简称Z分数）

Z标准分是将原始分数与其平均数之差除以标准差所得之商，即

$$Z=\frac{X_i-\overline{X}}{S}$$

式中：Z——标准分；

$X_i$——原始分；

$\overline{X}$——团体平均分；

S——团体标准差。

可见，Z分数是一个原始分数与其平均数的离差，以标准差为单位。它表明此原始分数离平均数的远近，即表明原始分数在平均数以上或以下几个标准差的位置，从而明确此原始分数在团体

中的相对地位。因此，把原始分数转换成 Z 分数的实质，就是把单位不等距和缺乏明确参照点的分数转换成以标准差为单位，以平均数为参照点的量表分数。

若 $Z=0$，则该生成绩正是测验团体的平均分（$\bar{X}$），处于该团体的中等程度；

若 $Z>0$，则该生成绩高于该团体的平均水平；

若 $Z<0$，则该生成绩低于该团体的平均水平。

Z 标准分具有下列性质：

（1）一组数据中各 Z 分数的平均数（$\bar{Z}$）为零。

（2）一组数据中各 Z 分数的标准差（$S_Z$）等于 1。

因此，它是一个以绝对零点做参照点和有相等单位的导出分数。

（3）由于标准分数是原始分数的线性转换，所以标准分数的分布形态和原始分数分布形状相似，因此它才能总是代表一组原始分数。

（4）如果原始分数呈正态或近似正态分布，则标准差取值范围大致从 $-3$ 个标准差到 $+3$ 个标准差。

注意：上面用线性转换导出的 Z 分数只有在分布形态相同或相近的情况下才能进行比较。如果两个分布的形态不同（如一个是正态，一个是偏态），那么相同的 Z 分数可能代表不同的百分等级，就难以准确比较了。

因此，原始分数转换为标准分的前提条件应是原始分数的分布为正态。如果是非正态分布的原始分数，则应对其作正态化处理，而后才可转换与比较。关于原始分数正态化处理的方法可参考有关教育统计学书籍。

（二）T 标准分

用 Z 标准分代表学生成绩虽然带来了各科成绩的可比性和可加性，但由于其数值常出现小数、负数，不易为人们接受，因而

教育统计中又常将 Z 分数转换成 T 分数。即把 Z 分数乘以一个常数 K，再加上一个常数 C：$T = KZ + C$。

常用的 T 分数有两种表示形式：

1. 百分制形式的 T 分数

这是以平均分为 50、Z 标准分乘以标准差加常数（1 或 2、3 等）表示的 T 分数。

其通式为：$T = (S+1)Z + \overline{X}$。

也常用：$T = 10Z + 50$。

2. 九百分制形式的 T 分数

这是以平均分为 500，常模标准差定为 100 来表示的 T 分数。其通式为：

$T = 100Z + 500$。若 $Z = \pm 4$，则 $T_{最高} = 900$，$T_{最低} = 100$。

这种 T 分数表示形式在国外标准化考试中已常采用，我国在广东高考标准化试验中用九百分制来统计和公布高考成绩（最高 900 分，容许至多 3 个学生可达到此成绩，最低分为 100 分）。经过实践，已为师生家长所接受。T 分数只有高低之分，但没有满分、及格或不及格之分。

（三）原始分数的正态化处理

原始分数转换为标准分数，如果考察原始分数在总体中的位置，则需要原始分数的分布为正态作条件，如果原始分数的分布呈严重偏态分布，则需要对原始分数进行正态处理，但它有一个前提，即所测量的心理特质的分布应该是正态，只是由于测验误差而造成实测分数分布是偏态时，才能进行正态化处理，下面以实例来说明正态化过程。（这里 $T_1 = 11Z' + 50$、$T_2 = 11Z' + 60$。）

小结：

（1）标准分在相对评价中有很重要的作用，但要注意它的使用条件，标准分是针对各种考试成绩的原始分数不具有可加性而提出来的一种相对的位置量数，它适用于各种考试或历届同科考

试成绩分布大致相同或相近的条件下，若成绩分布相差太大时（如标准差太大时等），要先对原始分数的分布进行正态化处理，然后才能使用标准分。

**表 8－19　原始分数正态化的过程**

| 原始分数 | 频数 | 标准分数 | 累积频数 (f) | 累积频数的中点 (cf 中) | 累积频数中点的百分位 | 转换标准分数 (Z′) | 分数 $T_1$ | 分数 $T_2$ |
|---|---|---|---|---|---|---|---|---|
| 92 | 1 | 1.45 | 30 | 29.5 | 0.983 | 2.12 | 73 | 83 |
| 91 | 2 | 1.35 | 29 | 28 | 0.933 | 1.50 | 67 | 77 |
| 90 | 2 | 1.26 | 27 | 26 | 0.867 | 1.11 | 62 | 72 |
| 85 | 1 | 0.78 | 25 | 24.5 | 0.817 | 0.90 | 60 | 70 |
| 84 | 1 | 0.69 | 24 | 23.5 | 0.783 | 0.78 | 59 | 69 |
| 82 | 3 | 0.50 | 23 | 21.5 | 0.717 | 0.57 | 56 | 66 |
| 80 | 2 | 0.30 | 20 | 19 | 0.633 | 0.34 | 54 | 64 |
| 77 | 3 | 0.02 | 18 | 16.5 | 0.550 | 0.13 | 51 | 61 |
| 76 | 2 | −0.08 | 15 | 14 | 0.467 | −0.08 | 39 | 59 |
| 75 | 3 | −0.17 | 13 | 11.5 | 0.383 | −0.30 | 47 | 57 |
| 74 | 2 | −0.27 | 10 | 9 | 0.300 | −0.52 | 44 | 54 |
| 72 | 1 | −0.46 | 8 | 7.5 | 0.250 | −0.67 | 43 | 53 |
| 71 | 1 | −0.55 | 7 | 6.5 | 0.217 | −0.78 | 41 | 51 |
| 70 | 3 | −0.65 | 6 | 4.5 | 0.150 | −1.04 | 39 | 49 |
| 66 | 1 | −1.03 | 3 | 2.5 | 0.083 | −1.38 | 35 | 45 |
| 60 | 1 | −1.60 | 2 | 1.5 | 0.050 | −1.64 | 32 | 42 |
| 40 | 1 | −3.50 | 1 | 0.5 | 0.017 | −2.12 | 23 | 37 |
| $\bar{X}=76.8$<br>$S=10.5$<br>$N=30$ |  | $Z=\frac{X_i-\bar{X}}{S}$ | 自下而上累加 | $\frac{C_{f上}+C_{f下}}{Z}$ | $\frac{C_{f中}}{N}$ | 查正态分布表 | $\bar{T}_1=50.0$<br>$S_1=10.8$ | $\bar{T}_2=60.0$<br>$S_2=10.8$ |

（2）标准分具有可加性，是指它比原始分数相加更为合理，但并非十全十美，在处理可加性问题时应与原始分数作比较，全面分析，不要把二者相互肯定或否定。

（3）同科同次考试，把原始分数转换成标准分没有多大意义。

（4）学校教学中经常性测验和毕业的水平测验是以教学大纲为参照系的考试，旨在评价学生达到目标的程度，宜使用绝对分数评价，标准分数起不了这个作用。

## 第四节　标准分在评价化学学业成绩中的应用

一个原始分数一经转换为 Z 分数，就可知道它在分布中的相对地位。当原始分数的分布是正态分布时，只要求出分布中某一原始分数的 Z 分数，就可以通过查正态分布表（一般教育统计书中均有此表）得知此原始分数的百分等级，从而知道在它之下的分数个数占全部分数个数的百分之几，进而明确此分数的相对地位。

### 一、在确定考生相对地位中的应用

例如，某生 Z 分数是 1，查表可知他的分数比平均分多一个标准差，比他分数低的考生占 84.13%，比他分数高的占 15.87%。若考生群体是 100 人，则该生在其中处于第 16 名的位置。

又例，某学科教师以百分制给测验评分，若规定 93 分以上为优，现全班平均分是 72，标准差是 9，问这一测验中得优级的概率有多大？换言之，班内得优级的人数占多大百分比？也即说得分大于 93 分的可能性有多大？

解答此题，首先要把 93 分转换成 Z 分数，再查正态分布表即可。

$P(X \geqslant 93) = P(Z \geqslant (93-72)/9) = P(Z \geqslant 2.33)$，查表得：$P = 0.99\%$

注意，有的书中表为：$Z = 2.33$ 时，$P = 49.01$，应该再作处理：

$P(Z \geqslant 2.33) = (50-49.01)/100 = 0.99\%$

## 二、在比较学生成绩优劣中的应用

### （一）一个学生同一学科不同次考试成绩的比较

见表8－20。

表8－20　一个学生同一学科不同次考试成绩的比较

| 测验次 | 原始分 | 班平均分 $\overline{X}$ | 标准差 S | 标准分 Z |
|---|---|---|---|---|
| 1 | 85 | 87.14 | 14.59 | -0.15 |
| 2 | 84 | 80.35 | 8.89 | 0.41 |

可见，该生两次考试成绩相差无几，但标准分却相差较大，一正一负，表明第二次考试成绩在班平均分之上，第一次却在其下。从表面上看，第二次比第一次分数低，实际却是保持了相对进步的势头。

### （二）一个学生不同学科成绩的比较

见表8－21。

表8－21　一个学生不同学科成绩的比较

| 学科 | 原始分数 $X_i$ | 班平均分 $\overline{X}$ | 标准差 S | 标准分 Z |
|---|---|---|---|---|
| 数学 | 65 | 69.35 | 12.67 | -0.34 |
| 化学 | 67 | 72.84 | 9.89 | -0.59 |

从原始分看，该生化学分高于数学分，但转换后，数学分在班级中的地位实际超过了化学分。

### （三）几个学生几门学科总成绩的比较

见表8－22。

从原始总分看，两个学生学习水平没有差别。但若以标准总分比较：0.91 >0.65，则29号学生比11号学生成绩好些。

表 8－22　几个学生几门学科总成绩的比较

| 学科 | 原始分数 | | X | S | Z | |
|---|---|---|---|---|---|---|
| | 11 号 | 29 号 | | | 11 号 | 29 号 |
| 数学 | 85 | 73 | 69.35 | 12.67 | 1.24 | 0.29 |
| 化学 | 67 | 79 | 72.84 | 9.89 | －0.59 | 0.62 |
| 合计 | 152 | 152 | | | 0.65 | 0.91 |

## 三、在高考录取中的应用

### （一）现行高考录取制度是怎样使用原始分划线录取的

现行高考录取制度是采用按原始分的总分划线录取的。并按各科的重要性和中学的学习时数规定了各科在总分中的权重，例如3＋2考试前，数学在理科类中占的比例 120/710＝16.9%，生物占 70/710＝9.8%等。这就带来了问题：如果各科试题难易程度不同（即体现为原始分的 $\overline{X}$ 高低不同），成绩分布范围不同（体现为 S 不同），则各科在总分中所占的比例实际无法实现。见表 8－23：

表 8－23　1985 年全国高考理科类考生各科成绩表

| | 政治 | 语文 | 数学 | 物理 | 化学 | 生物 | 外语 | 合计 |
|---|---|---|---|---|---|---|---|---|
| 满　分 | 100 | 120 | 120 | 100 | 100 | 70 | 100 | 710 |
| 平均分 | 68.1 | 56.7 | 60.0 | 54.3 | 60.2 | 33.2 | 61.6 | 394.6 |
| 实占总分 | 17.2 | 14.4 | 15.2 | 13.9 | 15.3 | 8.4 | 15.6 | 100 |
| 应占总分 | 14.1 | 16.9 | 16.9 | 14.1 | 14.1 | 9.8 | 14.1 | 100 |

从表中我们看到数学和语文本应各占总分的 16.9%，但实际只有 14.4% 和 15.2%，而政治应占 14.1% 却占了 17.2%，这说明当年高考总分划线时，政治成绩起了很大的作用，而语文、数学成绩的作用就相对降低了。

这种两个以上科目的考试，由于其团体分布的平均数不同，

标准差也不同，其原始分是不能比较的。这种原始分相加的方法就好象100港币加上100元人民币再加上100美元得到的“300”元一样，是不能反映其价值高低的。

（二）怎样建立高考标准分制度

1985年以来，广东省高考标准化考试改革开始了标准分制度的探索。最初是在英语科实现的，采用的是固定常模——即标准分 $Y=(X-M)/S\times12+60$，将原始分按公式转换成100分以内的标准分，这种转换在试题难度控制在0.6左右时转换出来的标准分和我国传统的分数制度相差不大，容易为社会接受。但仅在一科中作转换并不能体现出标准分制度的优越性，因此当时的目的主要是摸索标准分制度的可行性。

1987年后，广东开始对全省高考各科实行标准分转换，采用的方案与英语科85、86年的基本相同，但各科没有统一常模，如数学、语文平均分为72，生物采用42，其他各科 $\overline{X}=60$，标准差则有的12，有的15，有的取最靠近原始标准差的整数等。

由于录取是以总分来划线的，而总分没有转换标准分，所以总分还是不能体现出考生所在的位置，且常模不统一，累加所得的总分还不尽合理，这种标准分制度的不足之处表现在：

（1）由于试题不稳定及其他原因，考试后成绩不可能呈完全正态，若出现严重偏态，未经正态化转换的标准分，就不能很准确反映考生所处的位置。

（2）由于常模平均分多定在60，标准差定在15或12，在分数转换中凡超过100分的成绩都要被“拉”下来（以保证限定在100分内），这对尖子生不公平。

（3）由于限定在100分内，仍易使人误为标准分数仍是题目答对率的标准。

因此真正采用标准分制度，应将标准分作正态化处理，使高考分数直接体现出考生的水平位置，并最好是打破百分制的观念。

从1988年开始，广东省的高考标准分制度再上一个新台阶：各科采用统一的常模：平均分为500，标准差为100，使用正态化方式转换标准分；总分也采用相同的常模进行正态化转换。其步骤简述如下：

1. 高考后，将各科成绩原始分分别按百分位等级转换成正态化分数（Z分数），然后按公式 $Y=100Z+500$，将各科成绩都转换为以500为平均分，100为标准差的标准分，这就是向考生公布的各科高考成绩。

这就打破了以原始分100分为满分的界限，而根据考生水平和所处百分位位置给予较科学的评价，也改变了“分数等于答对率”的观念，利于接受标准分的新观念。

2. 将每个考生各科标准分乘上各科的权重（语文、数学为1.2，生物为0.7，其它各科为1），累加起来成为原始总分。

再将理科（包括体育）类和文科（包括外语）类考生的原始总分，分别再次正态化。即分别再按百分位等级转换成正态化分数（Z分数），再按公式 $Y=100Z+500$，将该正态化的总分成绩也转换为以500为平均分、100为标准差的标准分，此即向考生公布的高考综合分，省招办即以此划定录取分数线。

这种标准分制度的特点是：

（1）打破了原来以100分为满分的传统概念，有利于社会尽快接受标准分制度。

（2）使用正态化转换，使得各科分数以及总分都能够体现出考生在全省考生中的位置。

（3）各科及总分使用相同的常模，使用起来简单方便。

（4）各科标准分累加成总分时才乘上各科规定的权重系数，使得各科的权重得以真正实现。

（5）由于综合分是正态化后才转换成标准分的，所以综合分所代表的是确定的百分位，也就是说，只要知道了录取率，在高

考末进行之前就可以定出录取分数线，减少了每年公布录取分数线前对学校、考生及家长所造成的压力。考生拿到分数后也可以立即知道自己的分数是处于全体考生的什么位置上。

（三）怎样利于高考标准分在高考后甚至在高考前就确定录取分数线

例如：当知道了理科考生的总人数为86419人，又知道本科录取数是4278人，分数线按录取人数的120%划出，也即5134人，占总人数的5.94%，从标准分百分位表中即可查出录取分数为656分。见表8-24：

**表8-24　标准分与百分位对照量表（常模平均分500，S=100）**

| 标准分 | 在此分数以下考生百分比% | 标准分 | 在此分数以下考生百分比% | 标准分 | 在此分数以下考生百分比% | 标准分 | 在此分数以下考生百分比% |
|---|---|---|---|---|---|---|---|
| 310 | 2.9 | 460 | 34.5 | 610 | 86.4 | 760 | 99.53 |
| 320 | 3.6 | 470 | 38.2 | 620 | 88.5 | 770 | 99.65 |
| 330 | 4.5 | 480 | 42.1 | 630 | 90.3 | 780 | 99.74 |
| 340 | 5.5 | 490 | 46.0 | 640 | 91.9 | 790 | 99.81 |
| 350 | 6.7 | 500 | 50.0 | 650 | 93.3 | 800 | 99.87 |
| 360 | 8.1 | 510 | 54.0 | 660 | 94.5 | 810 | 99.90 |
| 370 | 9.7 | 520 | 57.9 | 670 | 95.5 | 820 | 99.92 |
| 380 | 11.5 | 530 | 61.8 | 680 | 96.4 | 830 | 99.95 |
| 390 | 13.6 | 540 | 65.5 | 690 | 97.1 | 840 | 99.96 |
| 400 | 15.9 | 550 | 69.2 | 700 | 97.7 | 850 | 99.976 |
| 410 | 18.4 | 560 | 72.6 | 710 | 98.2 | 860 | 99.984 |
| 420 | 21.2 | 570 | 75.8 | 720 | 98.6 | 870 | 99.989 |
| 430 | 24.2 | 580 | 78.8 | 730 | 98.9 | 880 | 99.993 |
| 440 | 27.4 | 590 | 81.6 | 740 | 99.18 | 890 | 99.995 |
| 450 | 30.8 | 600 | 84.1 | 750 | 99.38 | 900 | 99.997 |

又例理科考生56000人，录取5800人，按112%划线则应上线6500人，占总人数的6500/56000=11.6%，查上表88.5%（11.6%在其上）位置，知分数线可定在619（或620）分。这意

味着，在高考前，只要有了录取率，就可定出录取线。

（四）怎样用标准分反映考生在全体考生中的位置？

例如：某理科考生综合分为500，表示他处于全体考生50%的位置（表8－24亦可见）。1990年广东高考理科和体育类考生总数约78850人，所以该生大约是39425名。

又例：某考生外语成绩700分，表示他在全体考生中处于97.73%的位置（查表4－21知）。1990年广东全体考生数是127500人，则比他分高的约有2894人。

（五）怎样用标准分进行各科成绩之间的比较

例如：某考生数学700分，生物680分，查表4－21可知其数学成绩位于全省97.73%的考生之上，而生物在96.42%的位置。

又如：某校有50名考生，历史成绩平均为640分，地理成绩平均为580分，表示该校考生的高考历史成绩在全省91.9%的位置，而地理成绩在78.8%的位置。

再如：某校文科30名考生的综合平均分为540，理科40名考生的综合平均分是580，查知表示该校理科考生在全省中的水平（78.8%）要比文科考生的水平（65.5%）要高。

（六）原始分与标准分的相关情况

广东对1987年理科考生抽样5233人、文科1603人，作了原始分与正态化标准分转换后分数对比分析如下：

（1）从原始分与标准分的相关系数看，实行正态化标准分的转换是合理的——高度相关。

| 理科87年高考 | 政治 | 语文 | 数学 | 物理 | 化学 | 生物 | 英语 | 总分 |
|---|---|---|---|---|---|---|---|---|
| 相关系数 | 0.967 | 0.997 | 0.881 | 0.985 | 0.994 | 0.994 | 0.995 | 0.984 |

可知原始分的顺序与标准分的顺序几乎没什么变化。

（2）统计表明：理科5233人，按1987年原始分录取线513分

上线的有 683 人（13%），转换成标准分仍按同比例，标准分划线在 613 分，有 684 人上线。这样得出二个表：第一个是原始分上线（513 分以上）而转换成标准分后不上线（613 分以下）的考生成绩情况，有 47 人；第二个表是原始分不上线（513 分以下），转换成标准分后上线（613 分以上）的考生成绩情况，有 48 人（约占总上线人数的 7%），即 7% 的考生出现了差异。这也是正常的，若无差异就不必转换了。

## 四、在管理学生学习质量中的应用

标准分除了可对学生成绩作静态比较外，还能作动态性的比较。如根据标准分作出学习质量管理图，即可真实地反映出学生学习成绩是进步还是退步，这是标准分的又一重要用途。

### （一）学习质量管理图（$\overline{Z}$ 管理图）的制作

1. 把原始分数转换为标准分

以 4 号学生为例（表 8－25）。

**表 8－25　4 号学生的测验成绩**

| 测验次 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 原始分 | 75 | 88 | 90 | 77 | 75 | 84.5 | 80 | 76 |
| 班平均分 | 87.14 | 80.35 | 82.69 | 71.97 | 68.85 | 72.84 | 77.80 | 70.07 |
| 班平均差 | 14.59 | 8.89 | 13.00 | 12.03 | 12.97 | 9.89 | 10.56 | 12.51 |
| 标准分 | -0.83 | 0.86 | 0.56 | 0.42 | 0.47 | 1.18 | 0.21 | 0.47 |

2. 把标准分取平均值 Z

$$\overline{Z}=\frac{\sum Z}{m}$$

式中 Z——每次测验标准分；

　　m——测验次数。

上例　$\overline{Z}=[(-0.83)+0.86+0.56+0.42+0.47+1.18+0.21+0.47]/8=0.42$

以测验序号为横坐标，以 Z 为纵坐标作出各 Z 点，再以 $\overline{Z}$ 为标准作平行于 m 轴的直线，连各点成折线即可。

4 号学生学习成绩管理图如图 8－8。

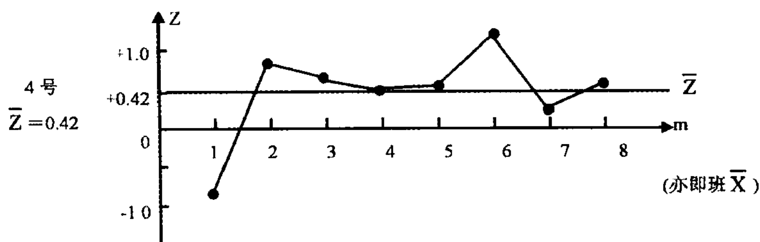



**图 8－8　4 号学生学习成绩管理图**

又例，24 号学生学习成绩（表 8－26）管理图如图 8－9。

**表 8－26　24 号学生的成绩**

| 测验次 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Z 分数 | 0.41 | －1.17 | 0.64 | 0.50 | 0.12 | －0.36 | －0.67 | －0.64 |

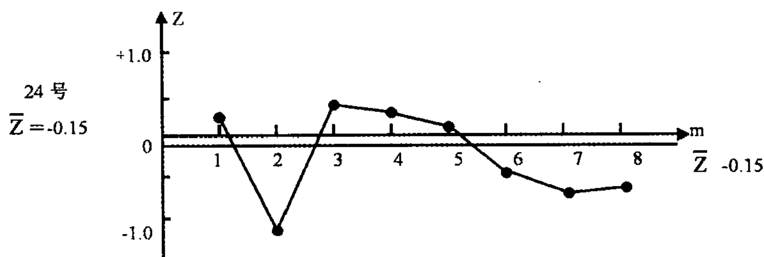



**图 8－9　24 号学生学习成绩管理图**

（二）$\overline{Z}$ 管理图的含义

（1）在某生的 Z 图中，若 $\overline{Z}=0$，表示该生成绩处于平均分的位置；若 $Z>0$ 表示该生成绩处于班平均分之上，如 4 号学生；若 $\overline{Z}<0$，则处于班平均分之下，如 24 号学生。$\overline{Z}$ 愈高愈好。

（2）当学生的学习状况无明显变化时，该生每次考试的标准

分在 $\bar{Z}$ 线上下分布，且落在上下的机会基本相同。

（3）有如下情况之一者，可认为学习有显著性变化。

1）连续有 5 个点偏在 $\bar{Z}$ 线一边；

2）连续 8 个点中有 7 个点偏在 $\bar{Z}$ 线一边；

3）连续 11 个点中有 9 个点偏在 $\bar{Z}$ 线 -边；

4）连续 13 个点中有 10 个点偏向 $\bar{Z}$ 线一边。

（三）$\bar{Z}$ 管理图的作用

（1）可以考察学生在集体中的位置。

（2）可以考察学生在某段时间学习质量变化情况。

（3）有助于教师了解教学效果。

（4）有助于把学习主动权交给学生。

由于质量管理图具有上述作用，可以真实地反映学生学习水平的升降情况，暴露出隐藏在数据背后的真实信息，揭示学生学习中的偏科现象，因此化学教师可以指导学生每人绘制 -份自己的学习质量管理图于卡片上，以便帮助学生自我发现问题。教师也可据此调节教学工作。

标准分在相对评价中有很重要的作用，但要注意它的使用条件。它适用于各科考试或历届同科考试成绩分布大致相同或相近的条件。对于同科同次考试，把原始分转换成标准分没有多大意义。学校教学中经常性测验和毕业的水平测验是以教学大纲为参照系的考试，旨在评价学生达到目标的程度，宜使用绝对分数评价，标准分则起不了这个作用。

## 第五节　相关统计与考试质量分析

在教育实践中，常常研究两种现象（两列变量）之间的变动关系，例如学习成绩与身体状况的关系，两门学科成绩之间的关系，高考成绩与大学学习的关系等等。两列变量之间的相互关系

在统计中简称相关，当一列变量增大（或减小）另一列变量相应增大（或减小），即变化方向相同，这种关系叫正相关，若一列变量增大（或减小），另一列变量变化方向与之相反，则称作负相关，若两列变量变化没有上述的相应关系则称零相关或不相关。

## 一、相关系数

表示上述相关程度的量叫相关系数，以 r 表示。

有了 r 这一统计量，我们就可以不但对两列变量变动方向是否一致进行研究，还可以对其一致性的程度进行数量化的研究。

r 被定义在 -1 到 +1 范围内取值，|r| =1 时为完全相关，|r| 越接近于 1，表明相关程度越高，r 值为 1 时，表示完全正相关。|r| 越接近于 0，表明相关程度越低；r 值为 0，表示零相关。相关系数的正负号只是表示相关的方向，其绝对值表示相关的程度，见表 8-27。

**表 8-27　相关系数与相关程度**

| 相关系数 | 相关程度 |
|---|---|
| 0 - | 互不相关 |
| 0.2 - | 略有关系，但不密切 |
| 0.4 - | 显著相关 |
| 0.7 - | 高度相关 |
| 0.9 | -极高度相关 |

相关系数的计算方法很多，其中最常用的是积差相关系数。

### （一）积差相关法

这是英国统计学家 Pearson 提出的，故被称皮尔森相关，其基本公式是：

$$r = \frac{\sum Z_X Z_Y}{N}$$

变换后又得到：

$$r=\frac{\sum(X-\overline{X})(Y-\overline{Y})}{NS_XS_Y}$$

其中 N 为成对变量数目，$S_X$ 为 X 变量的标准差，$S_Y$ 为 Y 变量的标准差，还可变换为

$$r=\frac{\sum(X-\overline{X})(Y-\overline{Y})}{\sqrt{\sum(X-X)^2\sum(Y-Y)^2}}$$

积差相关系数的使用条件是：

（1）二列变量确实相关。

（2）二列变量的分配形式都必须是正态的，积差相关系数适用于两个数列的变量都是连续变量，如百分制的成绩是连续变量，而五级分制就是间断变量。

（3）不要将不同样本容量的两个相关系数直接进行大小比较。

（4）一般地说，计算 r 的数据的对数 N 应大于 30，$N<30$ 时，计算出的 r 缺乏有效意义。

另外，两个变量 X 与 Y 的相关关系不一定是因果关系。如一群学生的数学与物理成绩的相关系数经计算为：$r=0.93$，这可以说明这些学生中数学分高的其物理分也相对高，两科成绩相关程度很高，但并不是说数学成绩好是由于物理成绩好，或反之，只能说两科有着很高的相关关系。

（二）等级相关法

若两个变量不是连续变量，或不是正态分布，或分布不清楚且为小样本，即对偶数目在 30 以下时，可用等级相关法。

等级相关法可适用于等级评定的情况，如思想优劣、身体好坏，或用来评价一份试卷的难易安排是否得当。

目前有不同的等级相关公式，这里只给出英国统计学家斯皮尔曼（Spearman）的等级相关公式：

$$r_R=1-\frac{6\sum D^2}{N(N^2-1)}$$

式中 D—两个变量之间每对数据等级的差数；

N—两个变量数据的对数。

（三）点二列相关

如果两列变量中，一列是正态连续变量，另一列属于只有两种分类的定性变量，如男和女，通过与没通过，重点与非重点等可化成“1”和“0”的变量，也称二分变量，用点二列相关就可以检验这个定性变量的两种分类对另一连续变量所产生的影响的一致性，其计算公式为：

$$r_{pq}=\frac{(\overline{X}_p-\overline{X}_q)}{S_X}\cdot\sqrt{pq}$$

式中 $\overline{X}_p$——p 部分的 X 数列的平均数；

$\overline{X}_q$——q 部分的 X 数列的平均数；

p——定性变量中某一类别的比例；

q——定性变量中另一类别的比例（q = 1 - p）；

$S_X$——全体 X 变量中的标准差。

点二列相关也是考试质量分析的重要方法，如鉴定试卷各题目是否有区分度，评价测验中某一问题与测验总成绩之间是否具有一致性等。如果一项考试规定：答对的记 1 分，答错的记 0 分，而学生考试成绩的总分数是连续变量，各题又是以错对分类的二分变量，这种真正的二分变量不能假设为正态分布。在这种情况下，就要用点二列相关的统计方法。

（四）用电子计算器计算相关系数

对于 n 值较小的情况，还可以用公式求算相关系数，但统计的样本较大时就麻烦了，通常可用电子计算器求之，但要注意一般计算器（如 CASIO 和 fx - 3600）中贮存的是积差法的计算程序，因此只能用于积差相关系数的计算，具体方法这里从略。

小结：不同相关的比较见表 8 - 28。

表 8-28　不同相关的比较

| | 符号 | 计算公式 | 适用范围 | 优点 | 缺点 |
|---|---|---|---|---|---|
| 积差相关 | $r$ | $r=\frac{\sum(X-\overline{X})(Y-\overline{Y})}{NS_XS_Y}$ | 两列变量均为正态分布的连续变量，且数据成对 | 求得的 r 精确可靠 | 使用范围受限制 |
| 等级相关 | $r_R$ | $r_R=1-\frac{6\sum D^2}{N(N^2-1)}$ | 两列变量均为等级资料，或可变为等级的资料 | 使用范围广，计算简单 | 分等易失信息，$r_R$ 不够精确 |
| 点二列相关 | $r_{pq}$ | $r_{pq}=\frac{\overline{X}_p-\overline{X}_q}{S_X}\cdot\sqrt{pq}$ | 一变量为连续变量，但另一列应为真正的二分变量 | 应用广泛 | $r_{pq}$ 值不够精确 |

## 二、相关系数的显著性检验

影响相关系数的因素除了一些偶然因素外，还和 N 的大小有关系。因此在求出相关系数后，还要对其真实性加以检验，这种检验称为相关系数的显著性检验。检验的结果使我们能在一定的可靠性水平上作出该相关系数的“不显著”、“显著”，或“非常显著”的判断。判断规则是：如果相关系数小于显著水平的临界值，则为不显著；如果相关系数等于或大于显著水平的临界值，则为显著；如果相关系数等于或大于非常显著水平的临界值，则为非常显著。

### （一）积差相关系数（r）的显著性检验方法

例如有 10 名学生的学习能力和学习水平的相关系数为 $r=0.78$，查《相关系数显著性临界值表》（一般教育统计学书中都有），df（自由度）$=N-2=10-2=8$，查表得：$r_{显著}=0.632$，$r_{非常显著}=0.765$，$0.78>0.765$，因此相关水平为非常显著。

（二）等级相关系数（$r_R$）的显著性检验方法

例如10名学生教育学与化学教学论课成绩等级相关系数为：$r_R = 0.87$，$df = N = 10$，查《等级相关系数显著性检验临界值表》得，$r_{R显著} = 0.648$，$r_{R非常显著} = 0.794$。

$0.87 > 0.794$，故相关水平为非常显著。即从总体上看，学生的教育学成绩与化学教学论成绩之间存在较高的正相关。

（三）点二列相关系数 $r_{pq}$ 的显著性检验方法和积差相关系数的检验相同

应当着重指出的是，相关系数只为我们指出了两列变量之间关系的性质和密切程度，其本身并不意味着因果问题。因果问题的判断还要依赖于在此基础上的进一步分析。

## 三、考试质量分析的基本指标

考试科学化的重要标志是对考试的质量进行数量化的分析，即通过一组数量指标对考试进行客观评价。那么，标志考试质量的数量指标都有哪些呢？从教育测量的角度对考试质量的分析主要包括整体试卷的分析指标，即信度和效度；各个试题的分析指标，即难度和区分度，一般将这“四度”作为考试质量分析的基本指标，下面分述。

（一）考试试卷质量分析的指标——信度和效度

1. 什么是信度

信度是指考试的可靠性，所谓可靠性，是指使用同一试卷测验多次，测量结果的一致性的程度，所以说，信度是对考试在测量上一致性的指标。

通俗地说，用一把尺子测量长度，多次测量，结果是一致的，我们就认为这把尺子是可靠的，具有高信度。

对于试卷的编制与考试，我们也希望有较高的信度，测量信度的基本方法是利用求相关系数的方法。

信度值的范围在 $0 \leq r \leq 1$ 之间，一般教学班级的测验要求信度值在0.6以上，大规模考试要求在0.9左右，最低不低于0.8，这样这次考试才是可靠的。

测量信度的方法很多，但任何一种表示信度的系数都是实际信度的一种估计。

（1）再测信度

以同一份试卷在不同时间去测验同一群体，求出两次考试分数的相关系数，就是再测信度，但是两次考试的不足是：考生的水平会发生变化，考生做过一次题目后，往往会记住了题目和答案，一般教师自编考试不宜用此法。

（2）“等值复本”信度

以两份等值（题型、题数、内容、难度和区分度一致）而又不同的试卷，在不同时间去考同一批学生，求得的相关系数就是“等值复本”信度，也称“平行试卷信度”。这种方法可避免考生机械重复前一次考试，但编制等值试卷是困难的，因此教师自编考试不大可能采用此法。

（3）分半法信度

为了克服上述由两次施测所造成的困难，可以在一次施测的情况下，通过某种手段将测试题目分成对等的两半，在记分时将此两部分的分数分开统计，计算其相关作为信度指标，称为分半信度系数。

为了使分开的两部分可以比较，常常是将一张试卷按奇数题与偶数题分半，如果题目是按由易到难的顺序排列的，那么奇偶分半法可得到近乎相等的两半。

使用分半法求得的两半分数的相关，只是半个测验的信度，全长的测验信度并不是两倍于一半测验的信度，但能用“斯皮尔曼－布朗（Spearman－Brown）公式矫正，若用 $r_{\mathrm{I\,II}}$ 代表一半测验的信度，则全长测验的信度为：

$$r_{信} \quad y = \frac{2r_{\mathrm{I}\,\mathrm{II}}}{1 + r_{\mathrm{I}\,\mathrm{II}}}$$

其中 $r_{\mathrm{I}\,\mathrm{II}}$ 可由积差相关公式求得，一般对考试信度多用此法求得：

（4）库德－理杰逊（Kuder－Riehardson）法：计算公式如下：

$$r = \left(\frac{N}{N-1}\right)\left(\frac{S_t^2 - \sum pq}{S_t^2}\right)$$

N—测验题的题目数；

p—每道题答对人数的比例；

q—每题答错人数的比例（$q = 1 - p$）；

$S_t^2$—整个测验的总方差（标准差的平方）。

此公式适用于测验的题型全部为选择题。

（5）影响信度的因素

影响因素很多，如试卷的长度是一个重要的因素，试卷长试题多，信度就高，所以加大考试份量是提高信度的方法之一。但若把信度 0.50 的试卷提高到信度 0.90，需要把考题增加到原来的 9 倍，实施是困难的（计算方法略）。所以要在试卷的编制上下功夫，如尽可能大题化小，多出选择题，既不超出时间及内容的限度，又可保证题目的最大数量，从而提高信度。

2. 什么是效度

所谓效度，是指一份试卷能够测量出它所要测量的东西的程度，它是考试的有效性指标。一个测验有时虽然能反映出受测者在某些方面的稳定水平，即有信度，但并不一定能反映出所要测量那个方面的水平，即没有效果。

作为测量工具的试卷是否有效，这与考试题目、考试内容都有很大的关系。考试效度的高低决定于它是否考了要考的目的和内容。例如，为了录用化学实验员，采用书面考试的形式而不考他的实验操作和动手能力，这样的考试的效度显然是不高的。

效度是程度上的概念，一个考试的效度可有高低之分，但不

能简单地说成有效与无效。

（1）效度的种类：常用的效度有以下三种：

①内容效度

这是指考试是否考了想考的内容的程度，一次考试不可能测量所有的教学目标，这就是测量目标恒小于教学目标。所以选择了适宜的内容，考试的内容效度就高，反之则低。

评价考试的内容效度，不是用统计的手段，而是根据教学大纲或考试大纲和命题的设计蓝图作出判断。

②预测效度

这里指考试是否起到了预测的作用，例如高校招生考试是否真的选拔的是该选拔的学生？这就要把学生的高考成绩和进高校后一年级有关学科的学习成绩进行比较，看看二者有无相关性。

预测效度用相关系数公式计算。如将考生成绩看作为 $X_1 \cdots X_n$，而将考生在大学一年级期末的成绩作为 $Y_1 \cdots Y_n$，计算其 $r_{XY}$ 的值。

③共时效度

共时效度和预测效度差不多，所不同的是比较两次同时举行的考试的成绩，这在建立一个标准化考试中尤为重要。因为我们对所建立的考试能否成为标准化考试，没有把握，具体做法是，让考生在差不多相同的时间内参加我们的考试和另一个公认是标准化的考试，然后计算两者的相关系数，如果相关系数高，则我们所建立的考试也应该是标准化的。

共时效度和预测效度都是把考试分数和一定的效标相比较，所不同的是，共时效度的效标是另一次标准化的考试，而预测效度的效标是考生将来的成绩。（注：人为确定的效度标准简称效标）。

（2）效度的计算

①在有可靠效标分数的情况下，用积差相关系数计算效度系数。

例如：某校高二进行的两次考试（假如第一次考试为效标，两次考试都符合标准试题要求，且内容分布和难度是同类和等值的），卷面统计成绩如下：

**表 8－29　学成绩的效度计算表**

| 学号 | 第一次测验 | 第二次测验 | $X=X_1-\bar{X}$ | $X=(X_1+\bar{X})^2$ | $Y=Y_1-\bar{Y}$ | $Y=(Y_1-\bar{Y})^2$ | XY | 计算 |
|---|---|---|---|---|---|---|---|---|
| 1 | 73 | 77 | +6 | 36 | +12 | 144 | 72 | |
| 2 | 71 | 80 | +4 | 16 | +15 | 225 | 60 | |
| 3 | 62 | 80 | －5 | 25 | +15 | 225 | －75 | $r_{XY}=\frac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum(X_i-\bar{X})^2\sum(Y_i-\bar{Y})^2}}$ |
| 4 | 69 | 66 | +2 | 4 | +1 | 1 | 2 | |
| 5 | 68 | 58 | +1 | 1 | －7 | 49 | －7 | |
| … | … | … | … | … | … | … | … | $=3425/4997$ |
| 47 | 66 | 71 | －1 | 1 | +6 | 36 | －6 | $=0.705$ |
| 平均分 | 66.83 | 65.17 | $\sum=65.5$ | $\sum=4301$ | $\sum=59.6$ | $\sum=5806$ | $\sum=3524$ | |

相关效度系数的价值还要通过查“相关系数显著性临界值表”去检验。上例，N＝47，df＝47－2，查表 $r_{显著}=0.288$，$r_{非常显著}=0.372$，$0.705>0.372$，$r_{XY}$具有极其显著相关性的意义。从而可作出判断，本次高二测试试卷的效度很好，第一次考试具有很强的预测力。

②在只有二分变量效标的情况下，用点二列相关系列计算效度系数.

例如：在某区进行了一次学科统考，希望知道这次统考的统计效度如何？

已知：被试者来自两所学校：甲校与乙校。甲校学生的学习质量高于乙校，我们以此为效标，采用点二列相关估计效度。计算如下：

总人数为400人，p为甲校被试百分比80/400=20%，q为乙校被试百分比320/400=80%，$\overline{X}p$为甲校成绩平均分：89分；$\overline{X}q$为乙校成绩平均分：81分；St为总体标准差：8。

代入公式：

$$r_{pq}=\frac{\overline{X}_p-\overline{X}_q}{S_t}\cdot\sqrt{pq}=\frac{89-81}{8}\sqrt{0.2\times0.8}=0.40$$

检验：df=400－2=398，查表得0.098、0.128，因为0.40＞0.128，所以甲校和乙校的差异显著。

由于效标效度采用相关系数表示，可以对被检测的测验进行定量化的分析比较，相对于内容效度而言更为客观，其意义也较直观，易于被理解和接受。但是，效标效度在实际学业测量中用得并不多，这主要是因为目前除高考外，还没有其他比较权威的学科考试，而高考却不能作为一般学科教学测验的效标。这是因为两者之间在内容、性质上都有很大的差别。高考是为高校录取新生服务的选拔性考试，而教学测验多是为教学服务的形成性测验。因此，寻找学业测验的合理效标，是学业测量研究的重要课题之一。

（3）效度指标对考试质量的衡量标准：见表8－30。

**表8－30　效度指标对考试质量的衡量标准**

| 效　度 | 1 | 0.7以上 | 0.4－0.7 | 0.3－0.4 | 0－0.3 | －1 |
|---|---|---|---|---|---|---|
| 有效程度 | 完全达到目的 | 优秀 | 良好 | 一般 | 较差 | 达到相反目的 |

（4）效度和信度的关系

信度很低的考试，效度不会很高；相反，信度高的考试，却不一定是高效度的考试。

效度系数≤$\sqrt{\text{信度系数}}$

小结：信度、效度的分类与计算方法可以归纳为表8－31、8

-32。

表 8-31 信度的种类与计算方法

| 类型 | 计算方法 | 适用范围 |
| --- | --- | --- |
| 再测信度 | 以同一份试卷在不同时间对同一群体重复测试，求出两次考试的相关系数。 | 异质性测验及速度测验教师自编测验难以采用 |
| 等值复本信度 | 以两次等值试卷在不同时间去测同一群体，求其相关系数。 | 难度测验、速度测验等。教师自编测验难以采用。 |
| 分半信度 | 实施一次测验，将所得分数分为等值的两半（如奇数题和偶数题），并算出每位受试者在两半测验的得分，最后再求两半测验的相关系数，并以“斯—布”公式校正。 | 同质性测验及难度测验，不适用于速度测验，因只须一次测验。教师自编测验常用。 |
| 库李信度 | 实施一次测验，计算每个题目的难度及每位受试者的得分，然后代入库—李公式：$r=\frac{n}{n-1}(1-\frac{\sum pq}{S_t^2})$ | 当试题为零壹记分法（选择题）时采用法。 |
| α信度 | 实施一次测验，采用柯氏公式：$r_\alpha=\frac{n}{n-1}(1-\frac{\sum S_i^2}{S_t^2})$，其中 n 为题数，$S_i^2$ 为每一题得分的方差，$S_t^2$ 为整个测验的总方差。 | 当测验中有主观题，又有客观题，出现非零壹记分时采用。 |

表 8－32　效度的分类与计算方法

| 类型 | 意义 | 计算方法 | 适用范围 |
|---|---|---|---|
| 内效度 | 考试内容对其所要测量的内容的代表性程度 | 目前尚无恰当的计算方法，主要采用逻辑分析方法，编制体现教学大纲的教学内容与目标的双向细目表，而后比较考试内容与此表的一致性，进而确定内容效度。 | 若以考试结果作为了解学生在某一学科的整体表现时，采用内容效度来评价。 |
| 效标关联效度 | 是指考试结果与作为“效度”的另一独立考试结果之间的相关程度。如果效标设置在考试之后，则为预测效度；如果考试分数和作为效度的分数两者获得时间间隔很短，就称为同时效度。 | 求考试成绩与其他考试成绩（效标分数）的相关系数，这里的其他考试成绩是用以估计本次测验的效度的标准，故称为“效标分数”。 | 若以考试成绩来估计目前或将来的成就在其他测验上的相关程度时，就用效标关联效度。 |

（二）试题质量分析的指标——区分度

对考试质量的分析与评价，除考查整个试卷的信度和效度外，还要对试题进行分析，教育测量学称这项工作为项目分析，从量的方面分析试题，主要是测量试题的难度和区分度。

1. 什么是难度

试题的难度是指试题的难易程度。通常用正确回答试题人数与参加测验总人数之比值作为指标，用 P 表示，其值在 0～1 之间，难度值越是接近 1，其难度就越小，正确回答试题的人数就越多，难度值越是接近 0，难度就越大，正确回答试题的人数就越少。这种顺序与习惯说法正好相反。

在编制试题时，人们希望选取难易适中的题目。如果一个试题的难度指标为 1 或 0，就无法区别受试者的知识水平和能力差

异，如果难度指标P越接近0.5，其试题将受试者的实际水平区别开来的效力越大。因此单从这个意义上讲，编制试卷应尽量选试题难度指数为0.5左右的题目。但是，由于都是P=0.5的题目过于同质，反而降低了总分的区别效力。所以在选题时，应在保证试卷平均难度接近0.5的前提下，注意选择不同难度的试题（一般在0.3~0.7之间选择）。大型考试试卷难度分布较好的应从易到难，逐步增加难度，且一般较易题约占25%，中难题约占50%，较难题或难题约占25%，试题的难度等级划分如表8-33：

**表8-33　考试难度等级的划分**

| 难度 | 0.9以上 | 0.9~0.7 | 0.7~0.4 | 0.4~0.2 | 0.2以下 |
|---|---|---|---|---|---|
| 等级 | 容易 | 较易 | 中难 | 较难 | 难 |

在编制试卷时，选用何种难度的试题，还应根据考试目的而定。比如，高考是选拔性考试，其难度值应视录取率而定，当录取率为20%时，难度指标P=0.2为宜；又如，试题为了检查给差生补课的效果，则一般将难度指标控制在0.7左右为宜。若要挑出10%的学生参加化学竞赛，试题难度应相对集中于0~0.2之间，使项目的平均难度控制在0.1左右。在标准化考试中，要求各试题难度控制在0.3-0.8之间。

2. 什么是区分度

区分度是指测验试题能否较好地将受试者的水平区分开来的量度指标，又称鉴别力。用D表示。区分度取值范围为$-1 \leqslant D \leqslant 1$，它对考试质量的评价标准如表8-34：

**表8-34　测验试题的区分度**

| 区分度指标D | 0.20以下 | 0.20—0.30 | 0.30—0.40 | 0.40以上 |
|---|---|---|---|---|
| 评价 | 劣等题 | 一般题 | 较好题 | 非常优秀题 |

3. 难度和区分度的关系

区分度与难度的关系密切，难度太大或太小，则所得结果可能是全体考生不通过或都通过，这样的题目就缺乏区分度。难度比较适中的试题，其区分度一般都达到优良水平。参加测验的学生数为100时，某试题的 $P=0.2$，说明有20个学生答对，80个学生答错，这时，学生之间相互比较的可能性为 $20\times80=1600$ 次；如果 $P=1$ 或 $P=0$，学生之间就没有相互比较的机会（$100\times0=0$ 或 $0\times100=0$）；而当 $P=0.5$ 时试题就能提供给学生相互比较的机会最大（$50\times50=2500$）。显然，试题的平均难度接近0.5，试题的鉴别力越强。

4. 难度和区分度的计算

这里将试题的难度和区分度的计算方法如表8－35。请注意使用条件，不可误用。

## 四、化学考试质量分析的基本方法

科学的化学考试质量分析方法，是借鉴教育测量学的理论，对化学考试试卷采用定性与定量相结合进行分析评价的过程。其中，在分析考生解答中的具体问题时，以定性分析为主，辅之以定量分析；在分析试卷和试题质量时，以定量分析为主，辅之以定性分析。

无论是定性分析还是定量分析，都离不开对试卷统计分析所得到的量化信息。因此，对试卷的统计分析是化学考试质量分析科学化的首要步骤和重要特征。

### （一）化学试卷的统计分析

化学试卷的统计分析的内容主要包括：考生成绩分布的统计分析、试卷质量的统计分析、试题质量的统计分析、考生答卷典型问题的统计分析。基本步骤如下：

表 8－35　难度和区分度计算方法一览表

| 类别 | 方法 | 公式 | 符号意义 | 适用范围 |
|---|---|---|---|---|
| 难度P | ①用通过率计算 | $P=\frac{R}{N}$ | R—该题通过人数，N -受试总人数 | 用于零壹记分法的客观题 |
| | ②用得分率计算 | $P=\frac{\bar{X}}{W}$ | X—考生该题平均得分，W—该题应得分。 | 用于得分不止两个确定的数值的试题，尤其是非客观题 |
| | ③用两端分组法 | $P=\frac{P_H+P_L}{2}$ | $P_H$（$P_L$）为高低分组答对该题的通过率或得分率 | 用于考生人数较多的测验，如年级以上测验 |
| 区分度D | ④用两端分组法 | $D=P_H-P_L$ | 同上③ | 同上③ |
| | ⑤点二列相关法 | $D=\frac{\bar{X}_p+\bar{X}_q}{S_t}\sqrt{pq}$ | $\bar{X}_p$—答对者测验总分的平均分，$\bar{X}_q$—答错者测验总分平均分，p—答对者占考生人数比例，q = 1 - p—（该题答错人数比例），$S_t$—所有被试测验总分的标准差 | 用于零壹记分题 |
| | ⑥积差相关法 | $r_{XY}=\frac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum(X_i-\bar{X})^2\sum(Y_i-\bar{Y})^2}}$ | X—每一考生某题得分，Y—每一考生全卷总分 | 同上②（即非零壹记分的试题） |
| 备注 | 两端分组法即将考生总分按从高到低的顺序排列，取前 27% 的人数为低分组。在求算 $P_H$、$P_L$ 时应注意按题目得分条件选用①式或②式。 | | | |

1. 抽样

对于较大规模的考试，一般都在全体试卷中随机抽样，用样本的统计分析来推断全体的情况。常用的抽样方法有：

（1）按系统抽样的方法，依序抽取每一考场首号考生的试卷

组成样本。如高考试卷分析即可采用此法。

（2）按考生在各个分数段的比例进行分层随机抽样。

①将全体考生的分数划分成若干个分数段。如每10分划为一段，即0～9分为一段，10～19分为一段，等等。

②统计出各分数段的试卷数量（考生数量）。如0～9分有8个；10～19分有15个；20～29分有33个，等等。

③计算出各段试卷数量占试卷总数的百分比。如0～9分的占0.39%；10～19分的占1.25%；20～29分的占3.87%，等等。

④根据确定的样本容量，按上述比例计算出各段应抽的样本数。如确定样本容量n=100，那么，按上述比例，0～9分抽1份；10～19分抽1份；20～29分抽3份，等等。

⑤用纯随机抽样法，按上一步骤计算出的份数，分别从各分数段中抽取试卷。为方便，可用随机数表（一般数理统计书中有此表）选取各段的试卷。

这种方法抽得的样本具有典型性，但必须作好总体分数段的频数统计，工作量较大，对于诸如高考类型的大规模考试不便采用。

（3）按学校类型的考生的比例进行分层整群抽样。

抽样样本n的确定方法有两种：一种是当总体很大时，就取n=370，这样便于用27%两端分组法统计分析（高考、会考等可以采用）。另一种通过计算来确定n，按统计上的要求，容量上千的总体，一般按1/30～1/20抽样，n可由此确定。

2. 考生成绩的统计分析

对考生成绩的统计分析首先是要整理样本分数，绘制频数分布表：

（1）按高分到低分排列成绩顺序，制作成绩频数分布图（图8－10）。

（2）计算样本特征量数平均分和标准差。

(3) 绘制样本分布直方图和频数分布曲线。

(4) 如果要与同科历次成绩比较，或与其他学科成绩比较，或与其他学科加和计总分排名次，还应将各考生的原始分数转换成标准分(T 分数)。若仅就同科同次考试而言，可以不作转换。频数分布图见图 8－10。

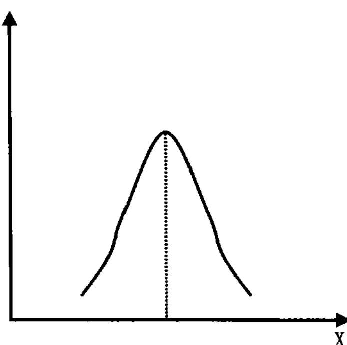



**图 8－10　成绩频数分布图**

再根据上述图表对考生成绩做出初步分析：

(1) 曲线呈正态分布，中间分数多，表明考生成绩正常，试题恰当，命题基本合理。

(2) 曲线高峰偏左（正偏态)，说明试题难度过大，或考生基础差

(3) 曲线高峰偏右（负偏态)，说明试题整体难度偏低，或基础好。

(4) 曲线呈驼峰，说明试题难度过大或过小，考生成绩异常，应从考生实际状况和考卷两方面找原因。

(5) 曲线呈陡峭型，说明考生基础相差很大或试题难度为中等的比例太大。

3. 试卷质量的统计分析

试卷质量的统计分析主要是通过对试卷结构统计分析的有关量表去体现，其中应将考试说明、教学大纲规定的有关比例、近年考卷有关的质量数据列入，以供对比。这些量表有：

(1) 试卷的知识结构。

(2) 试卷的知识内容在教材中的分布。

(3) 试卷能力结构。

(4) 试卷知识内容与能力要求双向细目表（据此表可以评价

试卷的内容效度）。

（5）试卷的题型结构和长度结构。

（6）试卷的难度结构。

（7）试卷的信度。

4. 试题质量的统计分析

为评价试题质量，主要制作如下统计量表：

（1）各题得分情况统计表（包括满分率、零分率、平均分、标准差、难度、区分度）。

（2）试题难度分布表。

（3）试题区分度分布表。

（4）选择题选项统计表。

（5）某些典型试题得分统计表。

然后再对试题的难度、区分度，选择题选项设计的合理性，选择题与主观题的分工与特色，各类型的特色等做出具体分析。

5. 考生答题典型问题的统计分析

为了全面评价考试质量，应注意搜集整理考生答题中的典型问题，并将其归纳成表格形式或其他便于直观的形式，通过数据分析进行解析。

以上统计分析多适用于较大规模考试，其中的统计计算多由计算机完成。对于常规考试或较小型的统考，若能借鉴其方法，将会大大提高考试质量分析的科学性。

（二）试卷的定性分析

有了上述统计分析资料，还必须通过定性分析的方法综合分析资料反映的问题和原因，对考试的质量做出质的评价。

定性分析的具体内容和方法应根据主考部门和考生所在单位的实际而定。如要了解考生答卷所反映出教学中存在的问题，可以根据统计分析提供的试题难度资料查找原因；若要了解不同地区、学校、班级考生的差异情况，可对考生的解答作比较分析；

若要了解命题的质量，应重点对考试的信度、效度，试卷的结构，试题的难度、区分度加以分析。

以上两大程序的相互配合，才能使试卷分析工作较为科学。到此，试卷分析的结果可用报告的形式写出，其中，可以侧重上述某一方面，也可以对上述诸方面加以综述，供有关部门及教学参考。

（三）试卷的讲评

试卷的讲评是试卷分析的一个组成部分。其对象主要是学生，把试卷定量、定性分析的结果告诉学生，使其真正了解自己的学习情况及相对位置。为此，试卷的讲评应包括：试卷的整体情况，答卷的整体情况，各试卷的具体情况等内容，并注意以下几点：

（1）针对性：即抓住典型错误，针对错误原因深入分析，有的放矢，因势利导。

（2）及时性：即利用学生考试后急于知晓结果的特殊心理，及时讲评，激发强化其求知欲。

（3）交流性：即安排学生之间交流，取长补短；开展教师之间相互切磋，教学相长。

（4）策励性：即对学习态度好而未考好的，适当批评指导，使其受到鞭策；对基础差而有进步的，热情表扬，使其受到鼓舞；对成绩一直优秀的，充分肯定，使之受到激励。

## 第六节　SPSS 软件在考试质量统计中的应用

利用电子计算机强大的数值运算能力，人们在化学教育科研中能进行庞大数量的数据处理，使过去许多束之高阁的统计分析方法得以为更多的化学教育研究工作者所使用，大大提高了化学教育科研工作的效率和结论的科学水平，从而促进了量化研究的开展。

社会科学统计软件包（SPSS——Statistical Package for Social Science）是教育与心理研究中最常用的商业化统计软件包，它的使用主要分两步：建立数据文件或数据库，执行数据分析。SPSS基本操作主要是输入数据和变量转换。

SPSS中的数据文件结构是二维的，纵列维研究变量，横列维是被研究的个体或被试。这种结构要求所有数据均需以数字编码的形式给出：名义测度（测度指在测量中被量化的程度）变量，如性别，男性编码定为1，女性编码定为2；序次测度变量，如受教育程度，文盲半文盲为1，小学为2，初中为3，高中为4，大学为5；等距测度变量，如摄氏温度；比率测度变量，如绝对温度。建立数据文件时，要求在同一列中对每一个被研究的个体或被试做出对应于变量的编码。

下面以SPSS for Windows为例简单介绍SPSS的基本操作，详细内容可咨询SPSS窗口中的Help。

## 一、SPSS统计软件的基本功能

### （一）SPSS for Windows的运行环境

SPSS是世界上最早采用图形菜单驱动界面的统计软件，它最突出的特点就是操作界面极为友好，输出结果美观、漂亮。它将几乎所有的功能都以统一、规范的界面展现出来，使用Windows的窗口方式展示各种管理和分析数据方法的功能，对话框展示出各种功能选择项。SPSS采用类似Excel表格的方式输入和管理数据，数据接口较为通用，能方便的从其他数据库中读入数据。

教育统计学中所采用的SPSS产品主要是SPSS for Windows。

### （二）SPSS for Windows的基本功能

SPSS共有三种显示窗口，分别为数据编辑窗口（Data Editor）、语法编辑窗口（Syntax Editor）和浏览窗口（Viewer）。

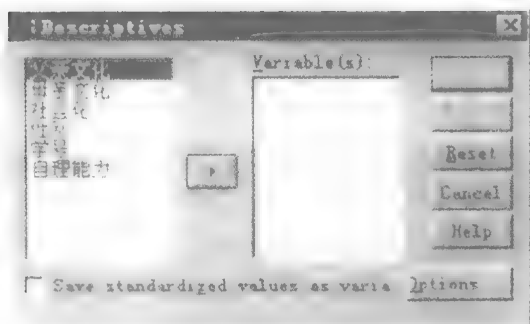
数据编辑窗口可以完成所有SPSS提供的统计分析功能。与其

他标准的 Windows 应用程序相似，SPSS 的数据编辑窗口包括标题栏、菜单栏、工具栏、状态栏，此外就是编辑栏、数据显示区（类似 Excel 的工作表区）。

## 二、SPSS 统计软件的基本操作方法

### （一）计算平均数的方法

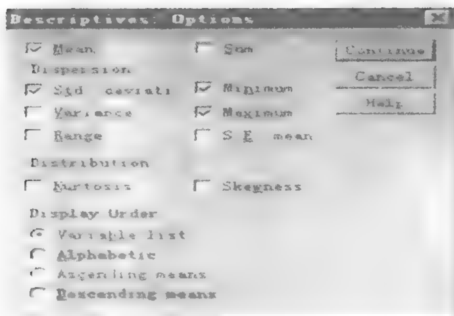
（1）启动 SPSS 程序，把保存的数据文件调入数据编辑窗口。




（2）单击 Statistics 菜单并把光标移到 Summarize 菜单项，此时弹出下拉菜单，将光标移至 Descriptives 菜单项，单击该菜单项，弹出对话框。

（3）把光标移至项目上并置亮它，然后单击对话框中间带箭头的按钮，项目便会移到对话框右侧的 Variable（s）下面的文本框中，此时 OK 按钮由灰变黑。

（4）单击对话框中的 Option，即弹出另一个对象框，选择要做的运算。

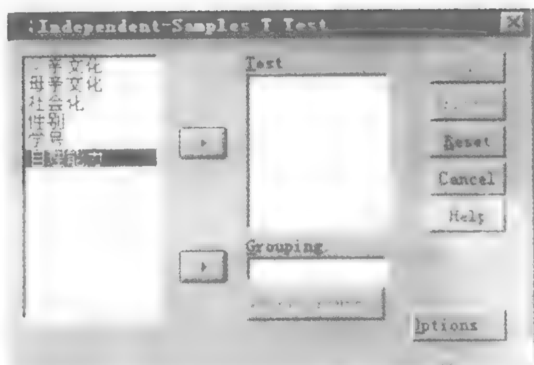


常用的选项有：

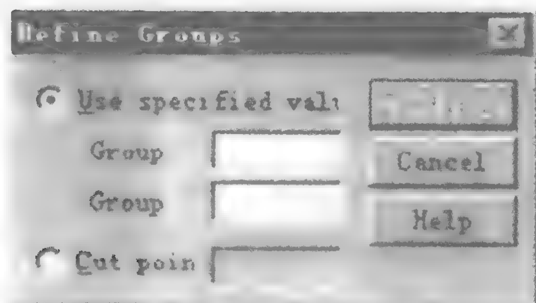
Mean：平均数；Sum：总和；Variance：方差；Minimum：最小值；Maximum：最大值。

（二）独立样本的 t 检验

（1）启动 SPSS 程序，把保存的数据文件调入数据编辑窗口。从 Statistics 菜单项开始按 Statistics——Compare Mean——Independent Sample T Test 顺序选择菜单，单击 Independent Sample T Test 菜单项，弹出对话框。



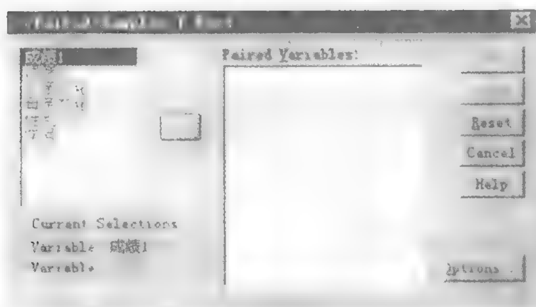

（2）把所要检验的项目选到 Test 下面的文本框中，把性别选到 Grouping 下面的文本框中。此时，Define Grouping 按钮由灰变黑，单击该按钮弹出对话框。



（3）界定分组变量，在上下 Group 后面的文本框中写上分组变量值，按 Continue—OK，输出结果。

（三）配对 t 检验

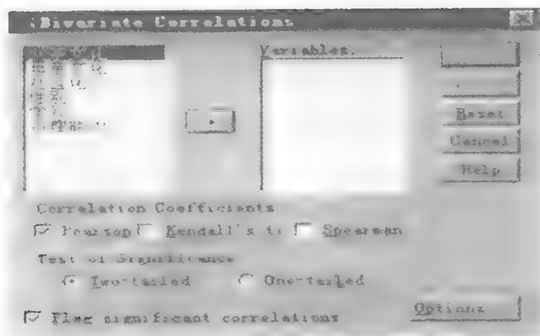
（1）启动 SPSS 程序，读入数据文件。从 Statistics 菜单项开始按 Statistics——Compare Means——Paried Sample T Test 顺序选择菜单，单击 Paried Sample T Test 菜单项，弹出对话框。




（2）把成绩 1 和成绩 2 选择到 Paired Variables 下面的文本框中，此时 OK 按钮变黑，单击，程序开始运行，结果输出。

（四）相关分析

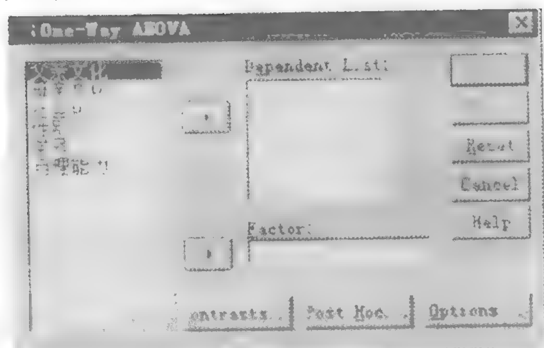
（1）启动 SPSS 程序，读入数据文件。从 Statistics 菜单项开始按 Statistics——Correlate——Bivariate 菜单项，单击 Bivariate 菜单项，弹出对话框。



（2）将所要分析的两个项目选择到对话框中的文本框中，此时 OK 按钮变黑，单击，程序开始运行，结果输出。

（五）方差分析

（1）启动 SPSS 程序，读入数据文件。点击 Statistics 菜单项，依次选取 Compare Means 和 One – Way ANOVA 项，单击 One – Way ANOVA 菜单项，弹出对话框。




（2）把项目选择到 Dependent List 下面的文本框里，另一项目选择到 Factor 下面的框里，此时单击 OK，系统执行方差分析运算，结果输出。

（六）$X^2$ 检验

（1）按 Statistics——Summarize——Crosstabs 的顺序选取菜单项目中的 Crosstabs 并单击菜单项。

（2）把选项选到 Row 下面的框里，把另一选项选到 Column 下面的框里。

（3）单击对话框下面的 Statistics，屏幕弹出对话框，选取 Chi – square，并单击 Continue，再单击 OK。

# 第九章　新课程学业考试改革

我国基础教育新课程改革的实施对学业考试改革也提出了新的要求。传统的学业考试的目标、内容、方式和评价方法等必须改革，才能适应新课程改革的发展。

## 第一节　学业考试改革的方向和目标

随着新课程改革的不断深入，学业考试的观念、考试内容、考试方式及成绩评定等存在的诸多问题也暴露出来。因此，研究我国学业考试的现状，认真分析学业考试仍存在的问题，从而确立学业考试改革的方向和目标，制定学业考试改革的策略，对学业考试进行合理地改革和完善，是十分必要的。

### 一、学业考试存在的问题

目前学业考试工作中存在的问题，可以从以下五个方面进行分析。

（一）*考试观念偏颇*

现行的学业考试，特别是一些较重要的考试如统考、期末考试、毕业考试等，常常成为评价学生学习效果的唯一方法，也成为评价教师教学效果和学校教育成果的唯一依据。学生考试的分数等级，常常作为区分学生优劣等级的依据，分数高就是好学生。教育行政部门也往往以统考成绩为依据来评价学校及教师工作质

量。另外，片面追求升学率的思想也使得学校被迫以考试为中心、为目的，把应该促使学生全面发展，旨在提高全民族文化素质的初、中等教育，变成了应试教育。这对培养人才是十分有害的。

（二）考试内容片面

目前学业考试着重考查学生对学科知识的掌握程度，而忽视了对学生智能发展的考查。试题编制偏重记忆、复述课本内容；有的试题内容偏离教学的重、难点，涵盖面狭窄，难以全面准确地考查学生的真实水平。

（三）考试设计落后

目前的学业成绩考试主要是经验型的，在试题的设计与检验上，很少进行科学的信度、效度检验；而且从出题，判分到评价，主观性较强。印象分、人情分常常影响考试的准确性。考试类型也比较单调，一般都是闭卷笔试，很难全面深入考查学生能力，特别是动手操作能力和创造力。

（四）成绩评定刻板

目前的学业成绩评定主要是依据所谓的“标准答案”，但有时一道试题很可能有多种解法，尤其是灵活题、论述题，学生完全可以发挥自己的创造力，从多种角度回答。而现行试题的评阅，则多按教师编制的一种答案给分，与答案一致则分高，与答案不一致则分低，这样就难以正确评定学生的真实水平。

（五）考试管理混乱

校内的学业考试，一般是由任课教师出题、做出答案并阅卷评分的。因此，任课教师出题后在组织复习时可能会出现漏题现象，造成“分数贬值”。另外，在考试的实施管理过程中，缺乏严格的组织措施，考场纪律不严，考生抄袭严重，也使得考试失效，不能准确考查学生的实际水平。

## 二、学业考试改革的方向和目标

针对上述学业考试中存在的问题，学业考试改革的方向和目

标应体现在以下几个方面。

(一) 更新考试观念

要正确认识学生学业考试的作用和意义，确立考试及考试分数的合理地位。既要认识到考试作为教学的一种手段，可以检验学生对知识技能的掌握程度，使教师与学生能及时调整自己的教与学的行为；又不能把考试当成教学的目的，过分夸大甚至滥用考试的功能。要科学分析考试成绩所反映出的学生知识掌握情况与智能发展程度，不为考试而考试，而是为更好地教学而考试。

(二) 拓展考试内容

考试内容的改革要体现新课程理念，全面反映各学科课程标准的核心要求。新课程背景下，学业考试不仅应对学生所学知识的掌握程度和解题的熟练程度进行检查，更重要的是对学生所应掌握知识的运用能力和相关技能进行有效的考核，特别是分析和解决生产和生活中实际问题的创造性能力。也就是说，考试内容应既重视基础知识，又突出综合能力，更强调创新水平。

(三) 严格考试环节

运用科学的考试理论指导实践，逐步使学业考试科学化现代化，包括从设计、命题、组织实施、评卷以至分数的使用一系列环节的科学化和考试手段的现代化。

(四) 完善评价制度

建立教师教学质量、学生学习质量的全面评价制度，完善学业考试的评定方法。

在计分方法上，不仅要有每个学生某张试卷的具体分数，而且要有能反映该学生在班、年级的相对水平及本人提高程度的分析，即要借助于教育统计知识，求得平均分数（$\overline{X}$），标准差（S），学科分数之间的相关程度（r），做定性和定量分析，才能正确反映某个学生的真实水平。

在成绩评定方面，可以根据实际情况采用具体有效的方法，

如累计综合评分法，即把该学生在某一学科的期末考试、平时分数及附加分数综合相加构成其总的学科成绩。这样就可以避免学生由于心理因素等引起的失误和误差。

（五）规范考试管理

中学学业考试改革中，必须进一步规范考试管理，建立科学和规范的学校考试管理体制及运行机制。应从以下几方面着手：(1）进一步建立完善、科学、严谨的考试制度，规范考试行为；(2）严格考试命题和把好命题关，确实做到教考分离；（3）严肃考试纪律，严格考场规则；（4）重视考试后的试卷评阅分析工作。

## 三、学业考试改革的策略

（一）变统一为选择

统一的考试成为阻碍学生个性发展的重要因素，变“统一”为“选择”，让学生根据自己的实际情况，选择最有利于自己发展的考试，能为学生个性发展提供宽广的平台，从而使学业考试由注重学生的横向比较，变为追求学生的纵向发展，成为促进学生不断超越自我的有利工具。如让学生从难易不同的试卷中选择试卷、选择试题和时间自主参加考试，也可以根据学科特点选择不同的方式参加考试。这种自主选择性的考试有效地避免了传统考试中纸上考实验、纸上考实践等形式主义的现象，全面地反映了学生的综合素质，促进了学生全面发展、和谐发展。

（二）变封闭为开放

封闭式考试限制了学生思维的发展，变封闭式考试为开放式考试才能使学生发散思维，见仁见智，各展特长。开放式考试要求考试形式的开放、试题本身的开放和试题答案的开放。考试形式的开放需要将传统的闭卷改为开卷或闭卷与开卷相结合的方式，也可以采用纸笔测试与其它考试方式相结合；开放性试题的目的是为学生提供广阔的思维空间，注重考查学生的潜能和素质，考

查学生能否灵活利用所学的知识，分析和解决具体情境下的具体问题；答案的开放性是学生在开放性试题的情境下，在已有知识的基础上，充分展示自己的个性，发散思维，自主选择问题解决方案，允许学生有个性化的表达，鼓励有创意的表达。

（三）变一元为多元

面向创新性人才的成长的学业考试改革，在制定考试目标和评价标准时，既要体现对学生的基本要求，也要关注学生个体的差异以及不同发展阶段的不同需求，发现和帮助学生发展其多方面的潜能，在评价标准、考试内容、评价主体、评价手段等方面体现多元化。评价标准的多元化就要考虑到不同学生的不同智慧类型，在保证其达到最基本要求的前提下，可以让不同的学生参加不同的测评；考试内容则须根据学科特色制定，从多个方面去考查评价学生；评价主体多元化是要让学生自觉参与到评价过程中，学会自我评价，成为自律学习者，同时让同伴、家长、老师共同参与到对学生的评价过程中，多方面多角度去测评学生。

## 第二节　学业考试内容与方式的改革

结合我国新课程改革和基础教育的实际情况，本节主要从学业考试改革的内容与方式两个方面进行论述。

### 一、学业考试内容改革

对学业考试内容的改革主要体现在科目设置、考查内容和试题命制三个方面。

（一）科目设置：统筹兼顾，合理引导中学教学

近年来，我国很多地区从减轻学生负担的角度出发，采取了减少升学考试科目的做法，并形成了以语文、数学、英语、物理、化学、政治等科目作为升学考试科目的基本格局。这种做法导致

的一个负面影响是：学校忽视甚至放弃非升学考试科目的教学，从而造成学生知识的结构性缺失。为避免这一现象，一些地区开始在升学考试科目的设置上做出调整。例如，在2003年的升学考试中，有59%的地区突破了按照传统六科设置考试科目的做法。其中，30%的命题单位采用不同形式将语文、数学、英语、物理、化学、政治、历史、地理、生物9个科目纳入初中毕业、升学考试中；14%的地区将地理和生物以外的7个学科纳入初中毕业、升学考试中；浙江各地的命题单位设置了语文、数学、英语、科学、社会5个科目。

根据《基础教育课程改革纲要（试行)》的要求，义务教育的初中阶段（7年级~9年级）设置分科与综合相结合的课程，包括思想品德、语文、数学、外语、科学（或物理、化学、生物）、历史与社会（或历史、地理）、体育、艺术（或音乐、美术）以及综合实践活动等科目。此外，为增强课程对地方、学校及学生的适应性，实行国家、地方、学校三级课程管理，鼓励地方补充开发适合当地需要的地方课程，鼓励学校补充开发或选用适合本校特点的课程。

2004年，在国家课改实验区中，有5个实验区将语文、英语、数学、物理、化学、生物、历史、地理、政治9个科目纳入学业考试科目，5个实验区将地理和生物以外的7个学科纳入学业考试科目，采用综合课程的5个实验区将语文、数学、英语、科学、历史与社会作为学业考试科目。此外，还有2个实验区采取了两考分离的组织方式，一个实验区将语文、数学、英语、物理、化学、政治和历史纳入升学考试科目，另一个实验区将语文、数学、英语、物理和化学纳入升学考试科目。

由此可见，为了确保在初中学校开齐、开足、开好国家规定的课程，促进学生全面发展，各地采取了增加学科考试科目的做法。

（二）考查内容：紧扣标准，全面反映学生的学习情况

本次课程改革实行国家基本要求指导下的教材多样化政策，鼓励有关机构、出版部门等依据国家课程标准组织编写中小学教材。因此，各实验区可以在多种不同版本的教材中进行选择。即使在同一个实验区，同一学科也存在不同版本的教材。显然，在学业考试考查内容方面，需要处理好课程标准与不同版本实验教材的关系。对于学业考试的考查内容，应以学科课程标准为依据，注重考查学生对重要知识与技能的掌握情况，特别是在具体情境中运用所学知识与技能分析和解决问题的能力。

各学科课程标准从知识与技能、过程与方法、情感态度与价值观三个维度，规定了该学科的课程目标（简称为“三维目标”），并清楚地描述了各方面的学习内容。然而，并不是学科课程标准规定的所有学习内容都必须纳入学业考试的考查内容。对于知识与技能，应重点考查学科中的核心知识、原理和应用技能，并符合课程标准规定的学习水平；对于过程与方法，应重点考查学生运用所学知识与技能分析和解决实际问题的能力以及探究新知识的能力；对于情感态度与价值观，不应采用直接设题赋分的方式进行考查，可适当地渗透在前两个方面的考查内容中。应该说明的是，没有试题能脱离特定的“知识与技能”而单纯地考查“过程与方法”和“情感态度与价值观”。

（三）试题命制：加强命题研究，提高试题质量

为了全面、准确地反映中学毕业学生在学科学习目标方面所达到的水平，需要进一步加强命题研究，提高试题的质量。学业考试的命题需要注意以下三个方面。

（1）充分挖掘客观性试题考查学生分析和解决问题能力的功能。

（2）有效发挥主观性试题考查学生的探究能力和实践能力的优势。

（3）加强试题与社会实际和学生生活的联系。

## 二、学业考试方式改革

考试的方式有很多，如笔试（包括开卷和闭卷）、口试、实验操作、撰写小论文等。但各种考试方式都不是十全十美，自有其优缺点，并且不可相互替代，这使得多种考试方法得以相继产生并同时存在。对于具体的一门课程，应采取怎样的考试方法进行学业考试，这需要按照专业和课程的特点，并结合考试的目的、考试的内容、命题的方式等因素做出科学的选择。对于同一门课程，为了全面考核学生的相关知识、能力与素质的需要，可将多种考试方法联合运用，采取分项考核、综合评定的办法，如实行开卷与闭卷相结合、笔试与口试相结合、理论考试与实际操作相结合等方法。其目的是为了科学、全面地鉴定学生的课程学业水平，促进学生对课程的全面掌握。

2001 年以来，课改实验区的学校在学期考试、单元测验、课堂测验等日常学业考试中，结合学科特点，积极探索了多种非纸笔的考试形式。在语文学科中，各实验区在纸笔考试的基础上增加了面试，通过朗读、背书、讲故事（说话）、演讲等多种形式，考查学生的口头表达技能。在英语学科中，增加了口语考试、听力考试等形式。一些学校还在英语考试中增加了课外表演内容，学生在课余搜集有关故事、诗歌、幽默、笑话等材料，借助字典、音像资料的帮助，准备一个 2～3 分钟的朗读、背诵或演讲节目，并可多次进行，由任课教师评价打分。大部分实验区在物理、化学、生物等学科中，增加了实验操作考试。一些学校将其分为书面测试题和平时成绩两部分。书面测试仍采用闭卷纸笔考试，设置了情感体验题和动手实践题两种新题型。平时成绩部分则由任课教师根据学生学习科学时的学习习惯、动手能力、小制作、作业等情况进行评分。一些实验区的历史与社会学科采用开卷考试。

学生在交期末试卷的同时，必须交上一份平时做的社会调查报告，调查报告以学生所在社区的人和事为题材，不交调查报告的，“考查通过社会调查获取有关知识”部分不能得分。一些实验区在音乐、美术等学科的考试中，让学生选送自己最满意、最喜欢的节目或作品展示。这些做法，较好体现了考试方式改革由单一性向多样性的转变。

## 第三节　多元化学业考试评价改革探讨

新课程倡导考试评价方式的多元化，以促进学生在知识与技能、过程与方法、情感态度与价值观等方面得到全面发展。这些多元化的学业评价方式主要包括纸笔测验、学习档案袋评价和活动表现评价等。

### 一、纸笔测验评价

纸笔测验是一种常用的重要而有效的评价方式，在新课程改革的推动下，纸笔测验也应根据课程改革的需要做出相应的改进。在中学教学中运用纸笔测验，重点应放在考查学生对基本概念、基本原理以及基本方法的相互关系的认识和理解上，而不宜放在对知识的记忆和重现上；应重视考查学生综合运用所学知识、技能和方法分析和解决问题的能力，而不单是强化解答习题的技能；应注意选择具有真实情景的综合性、开放性的问题，而不宜孤立地对基础知识和基本技能进行测试。例如下列问题就较好地体现了上述要求。

能源是人类生存和发展的重要支撑因素。请回答以下有关能源的问题：

(1) 我国的能源消耗以煤为主，煤炭的储量占世界储量的

13%，仅次于美国、前苏联，居世界第三位。据估计，全世界石油和天然气的资源于80年后将枯竭，煤炭最多供应200～300年。能源紧缺已成为世界经济和社会发展的一种制约因素，节约能源是解决能源紧缺的重要途径。根据你所学的化学知识，简述提高燃煤利用率的途径和方法。

（2）煤、液化石油气、天然气等能源以热的形式供给人们需要的能量。试以它们各自的主要成分C、$C_3H_8$和$CH_4$为代表，写出燃烧过程的化学方程式，指出它们在提供相同热量时哪一种燃料对环境造成的负面影响最小（燃烧1molC、$C_3H_8$和$CH_4$分别放出的热量为393.5kJ、2219.1kJ和890.3kJ）。

（3）传统化石燃料煤、石油、天然气的主要缺点是什么？为什么说“氢能源是最理想的能源”？

（4）在地球上，氢元素主要以水存在，请设计用水中的氢开发氢能源的方案。

新课程理念下的纸笔测验，实现在知识与技能、过程与方法和情感、态度与价值观三方面的综合评价。以化学学科为例，新课程的纸笔测验与传统纸笔测验相比，具体转变见下表9-1。

## 二、学习档案袋评价

### （一）档案袋评价的含义

档案袋（或称成长记录袋）评价是在20世纪80年代西方国家中小学评价改革运动中形成和发展起来的一种新的质性评价方式。它是指教师和学生有意识地将各种有关表现的材料搜集起来，并进行合理的分析与解释，以反映学生在学习与发展过程中努力、进步的状况或成就。档案袋可以说是记录了学生某一时期一系列的成长“故事”，是评价学生进步过程、努力程度、反省能力及其最终发展水平的理想方式。

表9-1 传统纸笔测验与新课程的纸笔测验的比较（化学）

| 项目 | 传统纸笔测验的侧重点 | 新课程纸笔测验的侧重点 |
|---|---|---|
| 知识 | 概念定义的背记、针对定义关键词的是非判断 | 利用概念来分析和说明有关事实和现象 |
| | 对每一个知识点进行孤立评价或者人为综合 | 对核心的重点内容进行整体考查 |
| 技能 | 化学计算追求形式和数字化 | 用反应事实化学概念和数学方法模拟解决真实的问题 |
| 探究能力 | 对实验操作的评价强调具体规范和细节 | 在实验活动中考查操作，重视操作的活动功能和目的 |
| | 实验室制取气体的具体步骤和注意事项的复述及是非判断 | 能够自己制取气体并进行简单的性质探究活动 |
| | 实验现象和化学方程式的记忆 | 生活现象和实验现象的说明和解释 |
| 题型 | 选择题、是非判断题所占比例较大 | 分析题、解释题、说明题、设计题将占相当比例 |
| 题量 | 题目数量很大 | 题目数量减少 |

学习档案评价是促进学生发展的一种有效评价方式。在档案袋评价中，应培养学生自主选择和收集学习档案内容的习惯，给他们表现自己学习进步的机会。学生在学习档案中可收录自己参加学习活动的重要资料，如实验设计方案、探究活动的过程记录、单元知识总结、疑难问题及其解答、有关的学习信息和资料、学习方法和策略的总结、自我评价和他人评价的结果等。教师应鼓励学生根据学习档案进行反省和自我评价，将学习档案评价与教学活动整合起来。

（二）档案袋评价设计的主要过程

虽然档案袋收集的材料是多种多样的，但它并不是材料的简单堆砌。如果缺乏良好的设计，档案袋就会变成单一的作品文件夹。本着科学有效、简单易行的原则，有关专家将设计与建立档

案袋的主要步骤分为以下三步。

1. 明确目的与用途。在设计档案袋的过程中，最为重要的就是“明确目的与用途”。档案袋的基本用途有三：第一个用途是展示，即用档案袋展示学生最好的作品；第二个用途是反映学生的进步，即通过形成性的评价，证明学生的进步；第三个用途是评价工具，即把档案袋作为一种总结性的评价工具。教师在设计档案袋的过程中，可以更多地将其作为反映“学生进步”和“展示学生作品”的工具。

2. 确定评价主体、评价对象及评价内容。档案袋十分注重学生在评价过程中学生的参与，如学生可以选择将什么装进档案袋，可以以多种形式表现自己的能力和水平。所以，在设计档案袋时，学生与教师一样被视为最重要的评价主体。此外，家长、管理者等也可以参与档案袋的评价。

3. 确定要收集的东西。档案袋中究竟应收集哪些材料呢？这取决于我们的评价目的。如果我们的目的是“展示”，那么只要收集学生最满意的作品即可。如果我们的目的是“反映学生的进步”，那么档案袋中既应收集过程性作业，也应收集结果性作业；既应收集学生的作业样本，也应收集其他一切可以描述学生进步的材料（如观察记录、他人的评价、测验试卷等）。同时，学生的自我反省和自我评价材料也可放入其中。

（三）档案袋的组成

一个完整的档案袋通常包括以下几个部分：

1. 封面：档案袋的封面应包括本人的姓名、年级、班级、学科等。

2. 前言：对有关学科学习档案的一个总的说明和提示，具体阐明档案袋中材料的种类、收集的方法等。

3. 目录：以提纲的形式展示各部分学习材料的联系，并标注页码。

4. 内容：包括核心内容和选择性内容。核心内容是每个学生档案中必须包含的内容，它为学生的自我评价和教师、家长及同伴的评价提供基础性的信息。选择性内容是学生自我选择的最能反映学生自己该学科学习情况的材料。档案袋的内容选择应由学生或由学生和教师共同讨论决定，一个完整学习档案袋中的材料可以包括：笔记、单元知识总结、疑难问题及其解答、探究活动的设计方案与过程记录、收集的学习信息和资料、学习方法和策略、课后学习感受、自我评价以及他人评价的结果等。凡是能够反映出学生在一定的时期关于学习过程中所做出的全部努力、进步、学业成就的材料都可以收录其中。需要说明的是对于每次收集在档案袋中的材料，都应注明日期，以显示随时间的推移所取得的进步情况。

5. 反思：对于自己学习的过程和结果方面所作的自省。如：我通过作业学到了什么；我哪些方面做得好；我为什么选择这件作品；在这个作业中我还要提高什么；我对我的表现感觉怎样；还存在哪些方面的问题。

6. 交流：档案袋中的材料虽然是学生自主构建的，但它却不是封闭的，因此，其中应显示同伴评价、家长评价和教师评价的一些信息。

（四）示例[①]

1. 化学学习档案袋设计

高中化学选修模块《化学反应原理》中“主题1　化学反应与能量”的学习档案袋：

（1）封面

（2）目录

① 杨香涛．新课程理念下化学学习档案袋评价的设计与实施探讨［J］．化学教育，2006，（5）．

表 9－2　“化学反应与能量”学习档案袋目录

| 序号 | 目录 | 页码 |
|---|---|---|
| 1 | 探究实验设计方案和实验观察记录（1）、（2） | |
| 2 | 调查报告 | |
| 3 | 资料（1）、（2）、（3）、（4） | |
| 4 | 学习化学反应与能量的有关知识时遇到的问题和教师对问题回答的记录 | |
| 5 | 当地收集到的常见化学电池标本 | |
| 6 | 历次试卷收集情况及自我订正与分析 | |
| 7 | 实验报告 | |
| 8 | 自我学习评价记录 | |
| 9 | 他人的评价记录 | |

（3）“化学反应与能量”学习档案袋中的材料

①有关化学能与热能、化学能与电能相互转化的探究活动资料（包括提出问题和假设、设计的方案、实验记录、对实验活动的自评和他评、对实验的反思和体会、问题讨论中的主要观点等）。

②对当地家庭使用煤气、液化石油气、煤等热能利用效率的调查和合理化建议。

③对反应热、焓变、盖斯定律、原电池、电解池、电化学腐蚀等内容的学习与认识材料，教师对问题回答的记录。

④收集到的与人类社会所面临的能源危机以及未来新型能源相关的资料；收集到的有关“化学暖炉”、“热敷袋”的构造和发热原理的资料；收集到的有关火箭推进剂的主要成分和燃烧热的资料；收集到的有关防止钢铁腐蚀方法的资料。如新闻、科技动态简报、图片、照片、实物等。

⑤调查当地市售化学电池，并展示收集到的电池标本进行交流、记录、提出的回收建议。

⑥学习化学反应与能量过程中历次试卷收集情况及自我订正

与分析记录。

⑦几种不同品牌电池工作效率比较的实验报告材料。

⑧对自己本单元学习状况的评价（如基础知识、实验探究和实验活动等情况）材料，以及有待改进的问题和改进的设想。

⑨小组同学的评价记录与任课教师针对本单元学习所写的激励性评语等评价材料。

2. 化学学习档案袋评价表

**表9-3 化学学习档案袋评价表**

《化学反应原理》主题1 化学反应与能量

年级　　班级　　姓名　　评价时间

| 评价内容 | 自我评价 | | 组内评价 | | 教师评价 | | 家长评价 |
|---|---|---|---|---|---|---|---|
| | 分数 | 质性描述 | 分数 | 质性描述 | 分数 | 质性描述 | 质性描述 |
| 探究实验方案设计和实验观察记录（2分） | | | | | | | |
| 调查报告和实验报告的撰写（1分） | | | | | | | |
| 学习资料和样本的收集（1分） | | | | | | | |
| 本主题涉及疑难问题及其解答（1分） | | | | | | | |
| 本主题知识测试、归纳和总结（2分） | | | | | | | |
| 试卷的订正与分析（1分） | | | | | | | |
| 与同学的交流和作品的展示（1分） | | | | | | | |
| 本主题学习的总结与反思（1分） | | | | | | | |
| 进步、个性或特色（1分） | | | | | | | |
| 总分、等级 | | | | | | | |
| 签名 | | | | | | | |

（注：等级判断　8分以上为优，7~8分为良，6~7分为中）

## 三、活动表现评价

### （一）活动表现评价的含义

活动表现评价是在学生完成一系列任务（如实验、辩论、调查、设计等）的过程中进行的。它通过观察、记录和分析学生在各项学习活动中的表现，对学生的参与意识、合作精神、实验操作技能、探究能力、分析问题的思路、知识的理解和应用水平以及表达交流技能等进行全方位地评价。评价结果以简单的方式加以记录，在比较、分析的基础上，给出恰当的反馈以激励学生进步。活动表现评价可以采用独立、小组或团体的形式；评价的对象可以是个人或团体；评价的方式可以是学生自评、同伴互评，也可以是教师、家长等对学生的评价；评价的内容既包括学生的活动过程又包括学生的活动结果；评价实施的时间可以在学习过程中，也可以在学习活动结束后。活动表现评价要有明确的评价目标，应体现综合性、实践性和开放性，力求在真实的活动情景和过程中对学生在知识技能、过程和方法、情感和态度与价值观等方面的进步与发展进行全面评价。

### （二）活动表现评价的特点

活动表现评价是一种不同于以往传统评价方法的新型评价方法。20 世纪 90 年代至今，学术界对活动表现评价的研究极为活跃，主要原因在于它可以评估学生的情感和社会技能，而不再局限于认知领域。这种评价要求学生实际完成某种任务或一系列的任务，如提出问题、做出假设、设计实验、做实验、操作仪器、辩论、调查、获得结论等，活动表现评价在活动中评价学生对知识与技能上的掌握，对科学探究过程的体验和感受，态度情感和价值观的形成，从活动中表现他们的综合成就，而不是从选择“ABCD”等答案来分析学生的掌握情况。

与传统的纸笔测验相比，活动表现评价的基本特点是：

1. 可以评价学生演示和运用概念与技能的能力、进行科学表征的能力。

2. 表现评价的任务涉及解释某个人的操作或所涉及的过程、设计的假设、解释科学条件、进行概括、描述各种形式或解法等。

3. 学生和教师事先已经知道有关要求，并且作好准备。事先向学生说明衡量标准，可让学生真正知道对自身的要求，并且让学生能主动地改造在技能或其他方面的弱点。

4. 教师将在特定的时间与状态下评价学生的表现，使学生不会感到意外。

5. 对学生的评价可针对他们是否能够适当与有效地应用以下项目：实验室仪器设备、测量工具、实验室安全程序等。

6. 操作是表现评价的一种方式。这种方式能有效地测知学生的理解能力，特别是配合面谈进行更为有效。

7. 评价通常通过表现任务的水平（难易、包容性范围等）及各分解任务的等级、任务对照表（对评定的内容的规定）或总体水平几个方面来进行。

（三）活动表现评价的操作流程

活动表现评价的实施可以按照以下具体流程进行，如图 9－1：

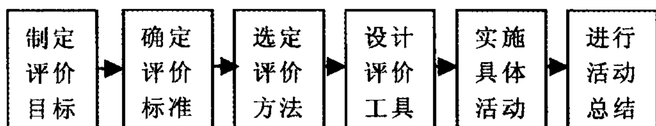



**图 9－1 活动表现评价操作流程图**

（四）活动表现评价的工具设计

1. 学生自评

在化学探究性实验活动表现评价中，可以采用表 9－4 让学生对照，进行自我评价或鉴定，将自己活动的感受、体验书写出来，让他们自己感受自身的进步。

表 9-4　学生自我评价表

活动名称　　　　学生姓名　　　　性别　　　　班级

| | 评价内容 | 质性描述 |
|---|---|---|
| 1 | 通过今天的这次活动，你对这种探究式实验感觉如何？你怎样看待它？ | |
| 2 | 动手操作前，是否有自己的猜想？是否计划好了操作的流程？ | |
| 3 | 活动过程中，你是否感到有困难之处？困难之处在哪里？你是如何解决的？ | |
| 4 | 你们是怎样相互支持和帮助的？<br>A. 分工合作　B. 1 人操作，1 人提供帮助　C. 先独立操作，再相互讨论<br>今后如果继续合作，你认为应该在哪些方面进行改进？ | |
| 5 | 活动过程中你停顿下来过吗？为什么停顿呢？ | |
| 6 | 对于实验中使用的仪器，你能够正确恰当地运用它们吗？你是如何用的？ | |
| 7 | 整个活动中，你是否向老师询问一些自己把握不住的问题？问题是什么？ | |
| 8 | 你认为这次活动效果怎么样？为什么？能不能再提出更多的问题？ | |
| 9 | 如果用等级评定（优、良、中等、合格和不合格）本次活动，你的自评结果如何？ | |

2. 小组互评

在活动表现评价中，学生小组成员之间可以用表 9-5 来进行小组互评。

表 9-5　学生小组互评表

活动名称　　　　学生小组　　　　班级

| | 评价标准 | 质性描述 |
|---|---|---|
| 1 | 活动前是否积极搜集有关资料？ | |
| 2 | 实验前是否有自己的假设？ | |
| 3 | 从哪儿搜集到哪些信息？ | |
| 4 | 在小组活动中完成了哪些任务？ | |
| 5 | 提出了哪些活动设想或者建议？ | |

| | 评价标准 | 质性描述 |
|---|---|---|
| 6 | 是否主动帮助他人解决问题？什么问题？ | |
| 7 | 主动配合他人完成了什么任务？ | |
| 8 | 认真倾听并采纳了他人的意见吗？ | |
| 9 | 与别人交流时表述是否清晰、简洁？ | |
| 10 | 实验结果如何？ | |
| 11 | 如果用等级评定（优、良、中等、合格和不合格）本次活动，评价结果如何？ | |

3. 教师评价

在化学探究性实验的活动表现评价中，教师评价常用的工具如表9-6，它可以对学生的活动表现进行较为客观的评价。

**表9-6　教师评价表**

活动名称　　　学生姓名　　　性别　　　班级

| | 评价内容 | 质性描述 | 评价等级 | | | | |
|---|---|---|---|---|---|---|---|
| | | | 优 | 良 | 中 | 合格 | 不合格 |
| 1 | 活动前参与对主题的选择 | | | | | | |
| 2 | 是否提出自己的假设 | | | | | | |
| 3 | 有没有合理的证据支持假设 | | | | | | |
| 4 | 设计的实验方案的合理性 | | | | | | |
| 5 | 根据实验适当的选择仪器 | | | | | | |
| 6 | 迅速正确的安装仪器 | | | | | | |
| 7 | 规范的称量药品 | | | | | | |
| 8 | 注意药品的节约 | | | | | | |
| 9 | 保持实验台的清洁 | | | | | | |
| 10 | 仪器的清洗 | | | | | | |
| 11 | 实验的结果 | | | | | | |
| 12 | 分析问题的思路是否清晰 | | | | | | |
| 13 | 与他人合作的态度 | | | | | | |
| 14 | 交流时语言表达的清晰度 | | | | | | |
| 15 | 参与活动的积极性 | | | | | | |
| 16 | 对于本次活动整体评价 | | | | | | |

（五）活动表现评价的成绩评定

活动表现评价是一种质性评价方法，它对学生在化学探究性实验中的种种表现和结果进行一种综合性评价，其对学生的情感态度和价值观这一项很难用定量的方法评价。因此，活动表现评价结果不能简单地只给学生打上分数或等级，而应该以质性描述为主。质性描述要使用激励性的语言来表达，构建一种有利于沟通和发展的心理氛围，以促进学生在其原有水平上进一步发展。但考虑到现行评价的实际情况和传统习惯，也可以辅以等级（分数）评定，如表9－4、表9－5和表9－6中包含有等级评定这个项目。在学生自评、小组互评和教师评价完成之后，再以一个比例做出综合的成绩评定，各种评价方法可以参照表9－7中的比例：

**表9－7　各种评价方法的比例**

| 评价方法 | 比例 |
| --- | --- |
| 学生自评 | 20% |
| 小组互评 | 30% |
| 教师评价 | 50% |

值得一提的是，在实施活动表现评价中还需要注意以下几个问题：

（1）活动表现评价中教师需要很多精力和时间，另外还要考虑财力、空间、时间和设备等方面的因素。因此，在选择活动表现的任务时一定要审慎。（2）在实施活动表现评价前要认真做好准备工作，评价结束后，要适时进行总结与反思，对评价活动中出现的失误及时补救、修正。（3）活动表现评价中，教师要面对很多学生，这可能带来主观随意性，影响评价结果的客观、公正和全面性。因此要注重活动记录的真实性、活动记录中应保留有学生的原始记录，以供确认。

以上介绍了三种学业评价方式。各种评价方式都有其优势与不足，学业成绩的评价不能只依赖其中某一种方式，而要综合使

用各种评价方式。

对学生学业成绩的评定结果可以采用定性报告（在写实性的记录基础上作分析性的描述）与等级记分相结合的方式呈现，充分肯定学生的进步与发展，帮助学生较全面地认识自己在群体中的相对水平，明确发展方向和需要克服的弱点。

另外，学业评价中还应注意：

(1) 注意知识的考评和能力、非智力因素的考评相结合。

(2) 注意平时的考评与期终考评相结合，注重形成性评价，建立系统的学业成绩评价体系。

(3) 注意书面考评与实验考评相结合。

(4) 注意学习质量的考评与教学过程的考评相结合。

(5) 注意绝对评价与相对评价相结合。

(6) 注意学生自评互评与他人评价相结合。

(7) 注意不要把考试成绩作为学业评价的惟一方法，倡导评价目标多元化和评价方式多元化。

总之，对学生的学业成绩和发展进行评价要注重评价与教学的协调统一，应建立包括纸笔测验评价、档案袋评价、活动表现评价等在内的多元评价系统，利用以学习成长档案为主的自我评价方法促进学生的学业进步与发展；利用活动表现评价学生的探究能力、学习态度、情感和价值观；利用纸笔测验评价学生掌握知识、解决实际问题的能力。努力改变过分追求评价的标准化和客观化的倾向，突出评价的整体性和综合性。

# 第十章　我国中考改革评述

初中毕业学业考试是初中毕业和高中招生制度改革的重要组成部分，它是“义务教育阶段的终结性考试，目的是全面、准确地反映初中毕业学生在学科学习目标方面所达到的水平。考试结果既是衡量学生是否达到毕业标准的主要依据，也是高中阶段学校招生的重要依据之一”①。这项考试的性质、任务和达成目标与以往的高中阶段学校招生考试都是不同的。因此有必要对初中毕业学业考试进行深入的实践探索和理论研究。

## 第一节　初中毕业学业考试制度改革

世界范围内，不同国家在中等教育第一阶段结束后的考试功能是各不相同的，归纳起来大致有两类：

一是单一功能，即考试的目的单一性。或是为了检查教育教学效果情况；或是作为毕业的依据；或是作为高中选拔学生的手段。

二是多重功能，指同一考试具有多种功能。即兼有毕业、升学、就业等多种资格的认定功能。通过学业考试，同时便获得了毕业资格、高中入学资格以及相应的就业资格等。

---

① 《教育部关于基础教育课程改革实验区初中毕业考试与普通高中招生制度改革的指导意见》，基教［2005］2号（下文简称《指导意见》）

根据这种划分方法，我国现行的“初中毕业生学业考试”，就应属于多重功能性质的考试，即兼有毕业、升学、就业等多种资格的认定功能。这和我们传统观念上的“中考”即初中毕业升学考试在意义上是不同的。

## 一、初中毕业生学业考试科目设置改革

为全面推进中小学教育教学改革，全面实现素质教育，《指导意见》指出：“学业考试应在课程内容结束后由当地教育行政部门统一组织。要减少考试科目，考试科目由省级教育行政部门确定或提出指导性意见。同时，学校还应对学生在综合实践活动等方面的学习情况进行考查，并体现在初中毕业生综合素质评价中”。

同时，为了能使初中毕业学业考试“全面”反映初中毕业生在学科学习目标方面所达到的水平，即考试科目设置应按照促进学生发展的目标体系，着眼基础性发展目标和学科学习目标两个方面进行全面考核，既要重视学生的学业成就，也要重视学生多方面潜能的发展，尤其是创新精神和实践能力的发展。评价标准应体现对学生的统一要求，同时要关注个体的差异以及发展的不同需求，为学生有个性的发展留有一定的空间。

如湖北省宜昌市的中考分为统一考试、考查与学校考查两种。统一考试的科目为语文、数学、英语、物理、化学；统一考查的科目为思想品德、历史、地理、生物、体育、信息技术技能和理化生实验操作；学校考查科目为音乐、美术、地方课程、校本课程和综合实践活动（研究性学习、社区服务、社会实践）；综合考查科目为体育与健康、音乐、美术、地方课程、校本课程和综合实践活动、语文听力与口语交际、英语听力与口语。

## 二、初中毕业生学业考试形式改革

根据《指导意见》的要求，“高中招生要坚持综合评价、择优

录取的原则。初中毕业生学业考试的方式应多样化”。

可实行校外考试（地方统一考试）与校内考试相结合；纸笔测验、听力测试、口试、实验操作相结合；闭卷、开卷或开闭卷相结合等多种不同形式。有些学科可以同时采用考核、考评和考查等多种方式。同时，还可以做这样的尝试“化整为零”的做法，将一次性中考化为初中各个学年段各课的考试。

校内考试考查的科目，一般应包括地方和学校编制的课程、综合实践活动、体育与健康、艺术（或音乐、美术）等。

安徽省2005年的中考改革方案中也是亮点频频闪现①：

1. 语文、外语、数学实行闭卷试卷，语文考试允许使用《新华字典》，外语考试加试听力；

2. 思想政治与历史（或历史与社会）、物理与化学实行开卷考试；

3. 从2006年开始，数学、物理与化学考试允许使用规定型号的计算器；

4. 物理、化学、生物实验操作的考试可由学生随机选择一项内容进行。

这些新举措、新变化都体现了新课程的理念。

### 三、初中毕业生学业考试成绩的新型呈现

《指导意见》规定：“提供给普通高中录取用的学业成绩应以等级的形式呈现，等级数和等级标准应由各地根据考试结果，并结合当地优质高中资源的实际情况确定。”

等级的划分不宜过多。一般可分为4～6等，能够满足三个或四个批次的录取需要即可。确定等级比例时，应主要考虑两个方

① 安徽省教育厅．安徽省基础教育课程改革实验区2005年初中毕业和高中招生制度改革方案，2005.

面，一是满足某一批次中一定比例（如85%～90%）的学生，能够在综合素质评定结果满足录取要求的情况下，以基本相同的成绩（如5A或4A1B），直接被相应的高中或中专学校录取；二是还应注意满足留出的10%～15%的名额，在适当扩大投档比例（如120%）后通过高中学校加试后录取。

等级划分时，可依照各学科的原始分数按一定比例划分，也可以将原始分转换为标准分后再进行等级划分，以满足不同学校的特殊需要，必要时（如特色学校招生）对某些学科的权重进行处理。

如安徽省2005年初中生学业考试成绩的呈现方式及运用：

（1）学业考试成绩以等级或等级与分数相结合的方式呈现。即条件成熟的实验区采用等级呈现成绩并用于录取；条件不成熟的实验区2005年暂时采用过渡的办法，即：等级附分数的办法呈现学业考试成绩，提倡和鼓励用等级作为录取依据。等级原则上分为A、B、C、D、E五个等级。

（2）体育、实验操作和学校考查科目的考试、考查结果必须体现在初中毕业综合素质评价中。其中物理、化学、生物实验操作以及综合实践活动的考查成绩可作为综合素质评价“学习态度与能力”维度的实证材料；体育考试成绩可作为综合素质评价“运动与健康”维度的实证材料；音乐、美术（或艺术）考查可作为综合素质评价“审美与表现”维度的实证材料。

## 第二节　初中毕业生综合素质评价制度改革

《指导意见》表明：“为全面反映初中毕业生的发展状况，应对初中毕业生综合素质进行评价，评价结果应作为衡量学生是否达到初中毕业标准和高中阶段学校招生标准的重要依据”。

本次课程改革的一个突破点在于：日常教学中的课程目标应

同时关注学生的知识与技能、过程与方法以及情感态度与价值观。因此，对初中毕业生的综合素质进行评价与新课程培养目标是相一致、相呼应的，这将成为推进课程改革、实现课程培养目标的重要举措。

## 一、初中毕业生综合素质评价的内容与方法

### （一）初中毕业生综合素质评价的内容

《指导意见》指出："综合素质评价的内容应以《通知》中提出的道德品质、公民素养、学习能力、交流与合作、运动与健康等六个方面的基础性发展目标为基本依据，各地可结合实际情况进行必要调整"。

### （二）初中毕业生综合素质评价的方式方法

为了确保综合素质评价的科学性和客观性，《指导意见》强调："实验区应组织专门力量制定综合素质评价方案。综合素质评价应以实证性材料和数据为基础，并充分考虑学生的自我评价和同学互评。评价时应注重对实证材料的分析与概括，力求做到客观与公正，避免以偏概全"。在具体实施中，综合素质的评价方式应注意做到以下几个方面：

1. 成立评定委员会

为了搞好初中毕业生综合素质评定，应在学校和区县（省市）成立两级评定委员会。

2. 制定初中生综合素质标准

为了确保初中生综合素质评价的客观性，需要制定明确的初中毕业生综合素质标准。对综合素质标准的各方面进行等级的划分，并对每一等级的关键表现及等级标准都作出详细说明。

3. 认真确定评价者

每个初中毕业班成立一个评价小组，由班主任和任科教师组成，也可以考虑学生同伴相互评价，但应充分认识到多种因素的

影响并权衡利弊。

4. 优化评价方法

对毕业生的综合素质评价以班级为单位进行。评价应以学生的日常表现为依据，通过观察、访谈和调阅学生的成长记录等方法给予学生客观和公正的整体评价。

## 二、初中毕业生综合素质评价结果的运用

### （一）各实验区的招生录取方案

目前在各实验区的招生录取方案中主要有四种使用方法：

1. 基础条件，是指录取时首先考虑综合素质的评定结果，规定考生在综合素质评定结果的最低标准（投档线或入围条件），而录取时主要按学业考试成绩进行择优录取。

2. 必要条件，是指在招生录取时，“综合素质评定结果”与“学业水平考试成绩”同时作为依据且具有同等重要的地位，两者不可偏废（如优质高中要求：学业成绩全优，综合素质评定结果亦需为优）。还必须对评定结果进行比例控制，即对学校的综合素质评定结果统一设置恰当的比例（如 A 等的比例略大于学业成绩的 A 等比例），防止综合素质评定的结果不等值或比例失当，而失去区分的意义或给招生造成困难。

3. 决定条件，是指录取时，首先考虑学业考试成绩；在成绩相同的情况下，按照综合素质评定结果来决定是否录取。

4. 结构条件，即作为录取的一项条件，相当于一门学科的成绩。

### （二）关于综合素质评定的等级划分[①]

划分综合素质评定的等级比例，亦应按照当地教育资源的状

① 贺斌．关于高中招生制度改革的思考与构想（下）［J］．教育理论与实践，2004，24（6）：27－31.

况及录取批次来确定。一般来讲，其 A、B 等级（或优、良等级）的划分，应比学业考试成绩的等级比例略大一些，以减少录取操作的难度。

综合素质评定一般由学校负责进行。为了体现招生录取的公平公正原则，保证评定结果的基本等值，各等级划分的比例，须由县以上教育部门根据招生涉及的地区统一规定。不同教育质量的学校和班级的比例可有适当的差别。

## 第三节　高中招生制度改革的探讨

高中学校招生除了依据学生参加初中学业水平考试的成绩外，将学生的平日表现和综合素质作为重要的参考和必备条件。

### 一、高中学校招生的基本原则

《指导意见》指出："普通高中招生要坚持综合评价、择优录取的原则。学业考试成绩和综合素质的评价结果应成为普通高中招生的主要依据"。

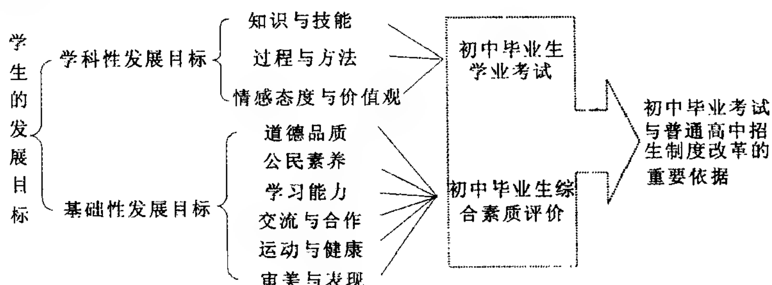



图 10－1　初中毕业考试与普通高中招生制度改革的重要依据

### 二、差额投档、分批录取

为了保证综合素质评定结果作为录取依据，同时为加试留出

空间，实验区及上级教育行政部门应根据招生计划，参照报考学生的学业考试成绩和综合素质评价结果，可按照“分批录取，留出适当空间；差额投档，加试研究能力”的办法进行招生录取。

所谓“分批录取”，一是根据教育资源状况划分批次，如可分为示范高中、一般高中、职业高中三个批次。相同的批次中，不同学校也还有计划指标、计划外调节指标或择校指标等。这些批次可以根据以往的录取办法因地制宜进行录取安排。二是指无论哪个批次的学校，都应采用“两批（次）录取”的办法。

第一次，依照学业成绩和综合素质评定结果首先录取招生计划的80%—90%（确定学业考试和综合素质等级比例时，应在总体上注意满足这一要求，但不一定满足每一所学校）。

第二次，利用剩余的指标采用“差额投档”。如投档的比例为剩余计划数的120%—150%则进行加试录取的办法，由高中学校根据学校情况组织专门委员会加试录取。

## 三、向高中学校下放招生录取的自主权

### （一）学生自由报名

学生从自己的实际水平和潜能出发，根据自愿原则，自由到招生学校报名。

### （二）下放部分考试权

各高中学校根据报考学生的学业考试成绩和综合素质评价结果，确定学校招生标准并进行录取，必要时可组织专门委员会加试。组织加试必须经地方教育行政部门批准。

### （三）学校自定录取方案

学校从本校实际和未来的发展出发，提前两年制定录取方案并向社会和初中学校公布，两年后可根据报名情况修订自己的方案。

## 四、多样化的高中招生制度的新尝试[①]

各地应按照有利于体现教育公平、均衡教育发展、扩大优质教育资源、促进学校创新发展的原则，结合当地实际，积极探索多样化的高中招生办法。

（一）综合评价录取制度

主要是指坚持将学业成绩和综合素质评定结果共同作为招生录取的主要依据，择优录取。这是2004年国家实验区招生制度改革的基本制度。

（二）招生名额分配制度

依据当地高中普及程度与各初中学校的办学水平，将部分招生名额直接分配到对口的初中学校，以缓解招生竞争给初中学校带来的压力。

（三）适当扶持薄弱初中制度

在对薄弱初中学校办学水平、教学质量进行考查的基础上，在本地区优质高中的招生计划中，划出一定的招生指标给相对薄弱的初中学校，或对薄弱初中的毕业生酌情降低招生标准。

（四）特殊学生推荐制度

对于品学兼优的毕业生或在某方面表现突出或具有突出特长的学生，以及平时表现和成绩突出，因故考试失常的学生，高中招生时可采用依据两名资深教师或学校推荐委员会的推荐，通过必要的调查核实或特长测评后破格录取。

（五）特色学校招生制度

鼓励学校根据自己的办学特色和人才培养需要，逐步扩大招生自主权。如外语特色学校，可以适当加试外语或进行口语测试；

① 贺斌．关于高中招生制度改革的思考与构思（上）［J］．教育理论与实践，2004，24（5）：21－25.

艺术体育特色学校可以增加艺术体育特长招生名额；优质高中应加强对学生科技创新特长、研究能力考查等，以逐步形成科技创新或研究特色。

（六）高中学校加试制度

要逐步扩大高中招生自主权，经地方教育行政部门的批准，高中学校可以组织加试，加试内容应主要考查学生的创新精神和实践能力，以及综合运用所学知识分析问题和解决问题的能力。

（七）免试保送制度

对于市地或县区级以上三好学生、优秀毕业生、优秀学生干部、优秀团干部等，经考查由所在初中学校保送直接进入高中学校学习。

## 第四节　新课程背景下中考命题特点与趋势

初中毕业考试的命题应根据学科课程标准并结合考试命题的指导思想，加强试题与社会实际和学生生活的联系，注重考查学生对知识与技能的掌握情况，特别是在具体情境中运用所学知识分析和解决问题的能力。

### 一、命题指导思想

初中毕业学业考试命题的指导思想是：

1. 有利于全面和准确地反映初中毕业生在学习目标方面所达到的水平。即要求初中毕业生掌握自身未来发展所需要的最基础的各种学科知识。

2. 有利于引导教师改变教学方式，促进学生主动、活泼、积极地学习。要坚持以学生发展为本，注重激发学生的学习兴趣，提高学生对知识和技能的掌握，促进学生学会解决问题的过程与方法，培养学生积极的情感态度与价值观。

3. 有利于普通高中在综合评价的基础上择优录取学生，促进普通高中的均衡发展。要求根据学生的发展水平进行命题，并根据义务教育各学科课程标准的要求，科学设置各试题，以保证不同层次的学生都能真实地表现出自己的水平。

## 二、命题特点与趋势

以“重视基础、关注探究、联系实际、促进发展”为基本理念的初中毕业学业考试命题将体现以下特点与趋势：

1. 根据课程标准确定考查要求

命题必须要遵循课程标准的要求，在考查学生基本知识和基本技能的同时，应重视考察学生的科学探究能力，同时体现课程标准对过程与方法和情感态度价值观的要求。

2. 遵循新课程理念设计试题

命题时凸显义务教育阶段各学科学习必须掌握的核心概念，关注学生对基本观念的形成与发展水平。

3. 始终关注对基础知识、基本技能的考查

基础知识和基本技能历来是教学与考试的主旋律，也是一个古老的课题，在改革中不能将这一奠基工程忽视。

但为了充分体现义务教育的性质，减轻学生过重的课业负担，避免教师在教学中对知识点死抠、深挖和押题现象，中考命题强调了依据课程标准和教材命题，充分考查基础知识和基本技能，不出偏、难、怪题，并保持适度的知识覆盖面。

4. 密切联系社会实际贴近学生生活

由传统的、单纯的关注知识转向以提高学生实际应用能力为导向，这就要求命题必须依据课程标准，以考查能力和素质为主，在继续加强基础知识考查的同时，着重考查能力、过程方法和情感态度价值观。不追求对所学知识的全面考查，而是突出重点，着重考查对学生高中学习乃至终生学习影响较大的知识和能力。

近几年各科的试题，都尽可能有效地摒弃了考查死记硬背的知识，试题的着眼点由知识转向了能力。试题大多从生产、生活实际中寻找挖掘命题素材。所要考查的知识都以生活中实际场景出现，目的是让学生用所学知识去解决实际问题，同时也引导教学走出死抠字眼、死钻教材知识的局面，教学中必须注重联系实际和提高学生的能力。

5. 增加开放性试题促进学生个性发展

由于试题强调了实践能力和创新思维的考查，各学科都出现大量不同形式的开放性、设计性和分析性试题。这类试题难度并不大，意在鼓励学生大胆想象、广泛联系，为学生提供了展示自己聪明才智的机会，有助于引导教师在教学中以学生发展为本，尽量发挥学生思维活跃的优势，为学生可持续发展打好基础。这对开启思维、激发创新，培养学生分析、解决实际问题的能力大有好处。

# 第十一章　我国高中毕业会考改革评述

高中毕业会考是国家承认的省级普通高中文化课毕业水平考试。它是检查、评价普通高中教学质量的一种手段，也是考核普通高中学生文化课学习是否达到必修课课程标准规定的基本要求的重要措施。然而近年来会考的弊端逐渐显现，会考的存废成为人们争议的热点问题。会考考什么？会考怎样考？会考究竟该何去何从？将是本章讨论的主要内容。

## 第一节　高中毕业会考的性质与功能

### 一、产生与发展

为了克服片面追求升学率的倾向，改变只把高考升学率作为评价高中教学质量的唯一标准的做法，教育部于1983年提出了高中毕业会考的创意。1985年全国第二届高考科研讨论会进行了专题研究，并提出建立会考制度，在会考的基础上进而改革高考，同年上海市率先进行高中会考；1988年浙江省首先取得在一个省的范围内进行会考的经验；1989年海南、湖南、云南三省实施会考；1990年国家教委颁发了《关于在普通高中实行会考制度的意见》，而后会考在全国逐步推开，到1993年全国大部分地区有30个省（市、区）都已实施。2000年起随着新课程方案的实施，会考管理权限下移，开始新的改革。至此，会考相继经历了酝酿、

创建、试验、局部推广、全面实施和转制等六个发展时期。实践证明，高中毕业会考制度有其存在价值，同时也需进一步完善。

### 二、性质与功能

实行会考的目的，就是从评价方面确立高中自身的目标，改变目前高中实际上把高考作为自己唯一目标的状态。具体而言，就是从我国普通高中的性质、任务出发，为学校确立一个比较科学、合理和权威的学科教学质量评价标准创造条件，通过正确的评价导向促进学校面向全体学生，全面贯彻教育方针和提高教育质量。

会考的性质是国家承认的省级普通高中文化课毕业水平考试，属于标准参照为主的终结性测验。它用于考核普通学生文化课学习是否达到必修课教学大纲规定的基本要求，也是检查、评估高中阶段学科课程教学质量的重要手段。会考于2000年起向社会开放，凡无高中学历的青年都可报考。

会考的功能是由其性质决定的。从现实和发展趋势看，会考已表现出来的功能有：教学检查、评价功能，少数高校选拔新生的依据，人才招聘的文化成绩依据等。会考最初设有11个科目，在高中三个年级分散实施。安徽省会考的科目自2000年起高一为地理、历史、信息技术，高二为数理化生（含实验考查），高三为政治、语文、外语、劳技。各门学科会考通过，即为合格的毕业生，方能取得该省的普通高中毕业证书。

## 第二节　会考的实施方法

### 一、时间与方式

会考一般安排在高一到高三分散举行。部分理科分考试（笔

试）和考查（操作）两种方式，由全省统一命题统一时间在指定的考场举行。操作考查亦为统一命题，一般由各地市或学校单独组织。考查时，考生通过抽签确定试卷中提供的题目中的某一题，而后动手独立操作，填写实验报告，教师当场检核评定成绩。

各省市、区教育主管部门在会考前要颁发《高中毕业会考考试纲要》。该纲要对会考的考核范围、水平要求、考核方式、考核目标、试卷结构等都作出了详细说明，并附有样卷。因而纲要是学生会考的依据，也是会考命题的依据。

## 二、试卷结构（以安徽省化学会考为例）

### （一）笔试试卷结构

1. 试卷组成

试卷由第一卷和第二卷组成。第一卷为选择题，一般每小题只有 1 个正确答案，共 30 小题，每小题 2 分，满分为 60 分。考生必须把自己认为正确的答案按要求填写在答题卡上。第二卷包括填空题、简答题和计算题三种题型，考生直接在试卷上作答，一般满分为 40 分。

2. 知识内容分布

基本概念和基础理论 22 ±5%，元素及其无机化合物知识 35 ±5%，有机化合物知识 15 ±5%，化学计算 16 ±5%，化学实验知识 12 ±5%。

3. 题型分布

选择题 60%，填空题和简答题 28%，计算题 12%。

4. 考试水平分布

知道 20 ±5%，识记 50 ±5%，理解 20 ±5%，应用 10 ±5%。

5. 难度分布

容易题 65 ±5%，中等难度题 25 ±5%，较难题 10 ±2%。

（二）操作考查试卷结构

1. 试题类型

按照常见的学生实验报告形式设计。

2. 考查内容分布

实验操作 70±10%，实验原理 20±10%，实验习惯 10%。

3. 考查水平分布

初步学会 50±10%，学会 50±10%。

（三）施测与评分

化学会考的施测和评卷由各地市组织进行，省级教育主管部门组织力量进行监测和抽查。分析和报告成绩采用等级分，有优、良、中、及格和不及格五级。

如浙江省评定成绩的办法是：首先将会考试卷原始分数与标准试卷量表分数等值，列出两个分数的对照表。然后根据对照表将每一考生的原始分数转换成量表分数。最后按下述量表分数区间划分等级并报告成绩，A 等：120 分到 96 分；B 等：95 分到 84 分；C 等：83 分到 73 分；D 等：72 分到 58 分，E 等：57 分以下。

## 第三节　会考的存废之争

### 一、废除会考制度的呼声

（一）高中毕业会考制度的弊端

高中毕业会考制度实践的过程中，逐渐暴露了它的一些弊端。

首先，实行“3＋X”高考以来，会考和高考分道扬镳成为高中教育的两个指挥棒。其次，高校扩招削弱了会考的必要性。再者，会考高度统一加重了边远地区农民子弟的经济负担。另外，会考已经滋长了腐败现象。

这些都在一定程度上加深了人们对会考的错误认识，学校、

教师和学生开始对会考的必要性产生怀疑，主张废除会考的声音逐渐浮出水面并占据绝对优势。一时间，“会考与高考背道而驰”、“会考占用学生大量高考复习时间”、“会考严重影响教师教学计划”及“会考滋生腐败”等等论说成了人们热议的话题。反对会考的声音越来越多，要求废止会考的呼声也越来越高。

（二）高中毕业会考改革现状

早在2000年，上海市教委就决定从当年开始不再举行全市性的高中统一会考，实施普通高中毕业考试制度，首先废除了会考制度。

2001年上半年，教育部出台《关于普通高中毕业会考制度改革的意见》，明确提出普通高中会考改革的统筹决策权将下放到各省、自治区、直辖市，一些素质教育搞得好的学校，可免会考，继续会考的地方，要适当减少科目，淡化会考成绩与高考报名的关系。之后，湖北、江苏等省做出了停止实施普通高中毕业会考制度的决定。陕西、山东、浙江等省对会考进行了改革，也部分取消了会考。到2006年为止，在全国29个省市会考取消与实施情况见表11－1：

**表11－1　29个省市会考取消与实施情况**

| | 省份 | 总数 | 所占比例 |
|---|---|---|---|
| 已取消会考的省（市） | 上海、湖北、贵州、山西、西藏、山东、广东 | 7 | 24.14% |
| 改革后仍在实行会考的省（市） | 北京、天津、重庆、湖南、江苏、河南、宁夏、黑龙江、新疆、辽宁、广西、浙江、吉林、四川、陕西、海南、河北、江西、福建、云南、甘肃、安徽 | 22 | 75.86% |
| 未做变动的省（市） | | 0 | 0 |

废除会考的省份越来越多，还在实行会考的某些地方，反对

的力量在高中教师的队伍中也占了绝对的上风。根据2003年《中国教育报》所做的一项网上调查显示，支持废除会考制度者占到了被调查人数的75%。

2004年，《中国教育报》发表的关于《高中会考制度，去还是留》话题讨论的读者来稿中，主张取消高中会考者认为——由于被地方操纵考试的“和尚”们搞得一团糟，高中会考正在成为最虚伪的一种考试；会考成了“全面提高学生素质”的幌子；会考+高考=招生；高中会考制度应该改革。而主张保留会考者认为——学生的全面发展是教育的首要任务；欠发达地区不能取消会考；会考需要加强管理；关键在于高考的改革。

2003年3月，一份《现行高中毕业会考制度弊大于利，应予以废止》的提案，出现在全国政协会议上。2005年6月，安徽省6位政协委员联名发出呼吁：“振兴教育，必须革新考试。会考这种有违初衷的虚伪、应付性考试应予以取消”。他们认为，会考正成为最虚伪的一种考试，这种流于形式的考试已经没有存在的必要，应予取消。

无独有偶，2004年，长春市12名人大代表呼吁取消普通高中会考，并在市人大会议上递交了“关于取消普通高中会考”的建议案。

## 二、会考制度必须存在

针对这些反对会考的意见，高中会考制度方案的论证和起草工作的参与者之一，华中师范大学考试科学研究中心的廖平胜教授则提出了截然不同的看法。

“会考制度存在的必要性和重要性要从高中教育阶段的重要性说起”。廖教授没有分析会考制度的表面现象，而是从理论的角度入手。“从一个人的成长来看，高中教育起到非常关键的作用。高中阶段在整个教育中的地位也是举足轻重的，是承上启下的阶段，

它的质量直接影响到高等教育的发展。如果一个人的高中知识体系是残缺不全的，这会使他在继续学习的过程中受到限制，而且，这种限制的形成将会因为人的成长而不可弥补。所以，高中毕业会考作为高中教育的终检环节，其产生是理所当然的。”

至于学生的负担问题，廖教授认为：“既是合理的，也是必要的”。首先，会考作为达标性的学业水平考试，它以国家规定的教学大纲的基本达标要求为依据而设计，其考查标准和内容，是一个合格高中毕业生必须掌握的内容和理应达到的基本目标。其次，会考这种负担，对高中生未来成才及我国教育水准的提升也是必须的、有益的。

以前，只有会考全部及格才能取得高考的资格，而现在的高考资格已经没有了这样的限制，江苏等省已经宣布，没有通过会考的考生也能参加高考。但是，这并不意味着会考失去了存在的价值。因为，高考是要看考生是否具备继续深造的能力或基础，而会考关心的是高中的知识掌握了没有。不能因为高考要选拔人才，就取消对高中教育的检测。考上大学只能说明他具备上大学的能力，而不能说明他就是一个合格的高中毕业生。高考无法起到监督教学质量和督促提高教育质量的双重作用。

现在，考上大学的高中毕业生只能占50%，如果没有了会考这个检测，将会有大量不合格高中毕业生得到高中毕业文凭，再加上高校招生规模的扩大，必然导致高校录取标准的下移，这样的结果将不能保证高等教育的质量。所以，必须发挥会考把关的导向职能。这方面的职能和功能，不能由于高校招生规模的扩大而消退。

廖教授指出：虽然现在的会考得不到尊重，没有得到应有的社会地位和价值，但是，无论从学生自身的全面发展和将来的发展还是从整个国家的教育科学化发展来看，会考都应该完善而绝不应该放弃。

另一种观点认为："会考制度非但不能取消，而且应该加强"。"水平考试和选拔考试合二为一是不现实的"。据吉林省普考办掌握的信息，1999 年到 2001 年，该省高中会考九科全 A 的人数分别仅占 1.9%、1.35% 和 1.6%，这说明要做到全优并不容易。

## 三、会考的问题探源①

会考制度从酝酿到推行至今已二十余年时间，然而制度还欠完善，存在不少尚待解决的理论和实践问题。

从高中毕业会考所负社会职能及设考目的角度考察，其现存问题突出反映在以下四方面：

### （一）功能局限

目前会考的功能主要体现为两点：一是检测评价高中教学质量；二是评定高考资格。其它方面的功能，如为人力资源开发、人才培养模式调整、师资结构优化、教育投资决策、教育结构调整等方面提供信息资源或参考依据，还未能落到实处，在很大程度上影响了会考的社会价值。

### （二）标准模糊

我国前期会考标准的确立，主要依据是高中教学大纲中有关教学要求的宏观性描述，很少有人进行学科教学与检测目标体系的研究，更无科学、规范并具操作性的学科标准体系，加之会考标准和内容的设计是分省进行，难免多种标准并存，故不利于各省成绩的比较和文凭的等值。

### （三）公信力差

由于会考仍处于健全完善之中，设计与实施尚缺乏科学理论的指导和严密的控制规程，以及会考宣传的欠广泛深入，社会对

---

① 廖平胜．论高中毕业会考的认识误区、社会价值与革新完善［J］．中国高校招生，2001，(1)：38－42.

会考的认同率较低，会考的信息资源尚未得到社会相关部门的应有重视和充分利用。

（四）权责不明

会考究竟由谁来管？组织机构如何设置？管理权责怎样划分？这些问题尚无明确具体的规定。现阶段，我国会考管理机构的设置十分零乱，这对会考职能的行使和功能的有效发挥极为不利，直接影响了会考权威的树立。

上述会考问题的存在，主要原因在于“三个不到位”，即认识不到位，条件不到位和管理不到位。

## 四、会考制度的革新完善

为使高中毕业会考制度能主动适应新世纪我国教育发展的需要，还必须切实抓好以下工作。

（一）重新定位

明确会考的地位和社会职能，是改革完善现行会考制度的立基点。在我国未来的教育发展中，会考至少应负起评价职能、诊断职能、监控职能、导向职能、整合职能等五大社会职能。

（二）重心下移

我国必须逐步将教育发展的焦点移向高中毕业会考科学性、规范性、权威性的强化，确保高中毕业生的质量。

（三）建立体系

会考的效率与效益要想达到我国基础教育发展、社会发展所期求的目标，需在现有基础上有计划地建立四大体系，即会考学科标准体系、会考质量评价体系、会考运作规则体系和会考社会服务体系，并大力开拓会考信息资源，为社会多方面提供信息和咨询服务。

（四）多元评价

随着高校和社会用人部门对会考信息服务需求的多元化，会

考检测评价领域也必须走向多元。一是进一步完善认知领域的评价；二是全力发展动作技能领域的评价，如计算机、书法、音乐、舞蹈、绘画、棋艺、科技等方面的技能水平评价；三是积极开拓情感领域的评价，如专业性向、能力倾向、责任感、事业心、意志力、协作精神、社会性等心理素质的测评，坚持为促进高中素质教育的发展服务，为高校及社会部门因材施教、因材施事服务。

（五）开架服务

现行会考科目的设置必须打破全国一统的僵化局面，理应根据实际需要有针对性的增减考试科目，灵活安排考试时间，实现会考的社会化。

（六）优化管理

主要解决三个问题：一是真正理顺权责关系，对于“谁来管”，“管什么”，“如何管”应作出明确的权利和责任划分；二是提高会考队伍的素质，通过各种有效的措施迅速提高现有会考人员的业务水平和管理技能，逐步达到专业化、职业化的标准；三是更新技术手段，不断提高对会考设计、实施的调控能力。

（七）加强研究

会考要发展，科研要先行。尤其是现行会考人员应增强科研意识，同时会考管理部门应组建一支专兼结合的会考科研队伍，对会考中存在的问题和面临的难题，开展有组织、有计划的研究，以此不断强化会考的理论基础，促使会考工作由经验型逐渐走向理性化。

## 五、部分省（市）会考改革的新方案

目前全国29个省（市），有7个省市已取消全省统一组织的会考。其中，贵州、山西、西藏三地彻底取消了会考，而山东省取而代之为学业水平考试；上海、湖北取代会考为高中毕业考试；广东省（全国首批4个高中新课程实验区之一）则采用学分制管

理，修满学分即可毕业。其余的22个省（市）都根据各省的实际对会考实行了或大或小的改革。归纳起来，目前实行的会考改革方案主要包括以下几点：

（一）会考改革的共性

1. 会考向社会开放，允许高中阶段各类学校（包括职中、中专、技校）的学生和社会青年参加，成绩合格者，颁发同等学历普通高中毕业证书。其中四川会考改革甚至允许初中生参加高中会考。

2. 实行每年1月和6月两次考试，允许学生跨年级参加会考，成绩合格者可提前毕业，提前参加高考；取消会考补考，允许学生重考，以最高一次的成绩为准。有的地方（如河南省）还规定每次考试开齐全部考试科目，考生可自主选择考试时间和科目。

3. 对符合条件的学生和学校试行全部免考和部分免考。

4. 会考的形式包括考试和考查两种，分为卷面笔试、实验操作、电脑操作等。其中广西省还包括体育测试。

5. 会考的内容包括学科知识、实验操作技能、信息技术基础、体育和劳动技术。在考核学科基础知识和基本技能的同时，注意考核能力。其中有的省市还逐步实现了信息技术考核“无纸化”，全部采用电子方式。

除宁夏自治区外，目前全国绝大部分保留会考制度的省市都在实行“信息技术”考核，主要考核学生的实际运用能力和动手操作能力。

（二）会考改革的特色

江苏将会考权利下放到各市、学校，9门文化课由各学校自行组织。

河南对政治试行开卷考；宁夏则实现了对政治、历史两科开卷考，另外增加了英语听力考试。

黑龙江规定语、数、外三科由学校自行组织考试；而陕西省

规定语、数、外、政治四科不再全省统一考，省示范高中可免考。

海南省实行高一不会考，文化科目的会考方式分为考试、考查两种，采用“3+3+4”的形式，即语文、数学、外语3门为必须考试的科目，在其余科目中学生可任选3门作为考试科目，余下4门作为考查科目（笔试）。

吉林省将创新意识计入会考成绩，平时成绩和统考成绩各占一半。

湖南在九科文化课会考的基础上增加“综合”一科，在高中三年级第二学期初开考。

浙江会考采取考试和考查相结合、统一组织和学校自行组织相结合的方式。文化课和专业基础课的考试科目为：语文、数学、外语和学生所学的1门专业基础课，由省统一组织。考查科目：文科班为信息技术、历史、地理、思想政治；理科班为信息技术、物理、化学、生物，由学校自行组织。

辽宁于2006年6月进行反向小综合会考。2006年起辽宁省实行反向小综合会考基础上的“3+小综合”高考科目设置方案。新方案中增加了反向小综合会考，即有报考文科倾向的学生，在高二期间需要参加理科综合的会考；报考理科倾向的学生，在高二期间需要参加文科综合的会考。其中，文科综合卷包括政治、历史、地理；理科综合卷包括物理、化学、生物。参加反向小综合会考的学生的成绩，可作为高校录取新生的参考成绩。

## 第四节　学业水平监测考试探讨

普通高中学业水平测试，又称为普通高中学业水平监测考试，或普通高中学生学业基础会考，前身即普通高中毕业会考。学业水平监测考试是在总结已经实施了十几年的高中毕业会考制度基本经验的基础上，为适应新课程改革的需要而实施的一项重要制

度，它将取代实行数年的会考制度，成为兼顾高中毕业与选拔功能的新考试制度。

## 一、学业水平监测考试的发展

2004 年 9 月，山东、广东、海南、宁夏四省区首先进行普通高中新课程实验，2006 年宁夏自治区、山东省与海南省率先实施新课改下的普通高中学业水平考试（基础会考），其中宁夏和山东对学业水平考试的成绩进行了 A、B、C、D 分级，而海南省将高中阶段参加的基础会考成绩折合成 10% 计入高考成绩，在改变“一次考试定终身”方面做出首个尝试。2007 年上海市、江苏省和福建省加入学业水平测试（基础会考）方案实施。其中江苏省方案里对选修科目的设置引起激烈而广泛的争议。2008 年安徽省紧跟其后，开展课改后的学业水平测试。2009 年广东省和黑龙江省也将先后加入课改的新一轮评价制度实施。

## 二、学业水平监测考试的指导思想

实行普通高中学生学业水平统一监测考试的指导思想是为了建立普通高中教育质量监督检查机制，规范普通高中的教育教学行为，使普通高中按照国家新的课程方案和课程标准组织和开展教育教学活动，全面提高教育质量。实行普通高中学生学业水平统一监测考试，有利于促进普通高中的课程建设和师资队伍建设，有利于促进普通高中办学条件的改善和优化，有利于促进普通高中的均衡发展，有利于推进普通高校招生考试制度的改革。

## 三、学业水平监测考试的性质

普通高中学生学业水平统一监测考试是普通高中课程改革的重要组成部分，是国家承认的省级普通高中学生文化课毕业水平考试。它是检查、评价普通高中学校教学质量和学校学分认定工

作质量的重要手段，也是考核普通高中学生文化课学习是否达到课程标准规定的毕业要求的重要依据。学业水平监测考试，是考试与学生评价制度改革的重要环节，是普通高中综合评价的重要组成部分，是依据国家统一标准，在国家指导下由各省份组织实施的全面检测普通高中学生课程修习水平的标准参照性考试。

## 四、学业水平监测考试的功能

为普通高中教育教学质量进行管理和监控，规范教学行为，确保学分管理质量，为师生的教与学提供指导；检测学生学科基础目标达成度和知识体系形成状况，保证学生全面素养形成必备的科学文化基础；为高校招生选拔提供参考依据；建立普通高中教育教学质量检测机制、规范普通高中的教育教学行为，使普通高中按照国家课程方案和标准开展教育教学活动，督促检查普通高中学校开全课程、开足课时，促进学生全面、主动、健康发展，具有重要意义。开展学业水平检测考试还可以解决以下问题：

### （一）检测高中教学情况，为评估学校教育、教学绩效提供重要依据

普通高中教育是基础教育的最高阶段，其价值取向和办学水平直接关系到人才培养的质量，政府必须把握和控制基础教育的人才培养目标、标准和方向，对教育教学质量进行有效的监控，才能保证其达到基本办学质量。建立高中学业水平考试还可以淡化实际存在着的高考对中学教学的评价作用。

### （二）兼顾毕业与升学的双重需要

高中学业水平测试既有通过考试的要求，同时又与升学挂钩，兼顾了高校招生选拔的功能，由于全部10门科目的成绩都将影响考生的升学情况，因此可以避免学生偏科。由于考试的成绩贯穿高中三年的学业，作为一个非常系统的考试体系会对高中学生做出一个全面客观的公正评估，这种考试评估体系对于高考的一系

列优惠政策更能体现出公平公正的原则，而不是一考定终身。

（三）掌握新课程实施情况，为新课程的顺利实施提供正确导向

高中新课程标准刚刚开始实施，尚在探索、建设、完善过程中，通过建立高中学业水平考试可以掌握各学校实施新课程标准的情况，为新课程的顺利实施发挥正确的导向作用，并确保学校开足所规定的科目和模块。

（四）促进学生对所学课程进行整合，构建知识网络

高中学业水平考试是按照科目进行的，这样可以使学生比较系统、全面地对所学的课程模块进行整合，建立相关模块之间的有机联系，构建知识网络。同时，可以根据统一的标准检测学生修业状况，为评价其是否达到课程基本标准提供依据。

## 五、学业水平监测考试方案简介（以安徽省为例①）

（一）科目设置

考试科目：语文、数学、外语、思想政治、历史、地理、物理、化学、生物、信息技术。

实验考查科目：物理、化学、生物、通用技术。

（二）考试时间与分值

1. 学业水平监测考试每年组织2次，时间安排在每年的1月初和6月中旬，每次开考10个科目。学生必须获得某科目所有必修模块的学分后，才能申请参加该科目的学业水平监测考试；学校必须在学生修习完成某科目所有必修模块，并进行了该科目的学分认定的基础上，才能组织学生参加该科目的学业水平监测考试。在此前提下，学生可自主选择参加学业水平监测考试的时间。

2. 学生在校学习期间可报考两次，以成绩最好的计分。成绩

① 安徽省普通高中学生学业水平统一监测考试方案，2008.

不合格的允许参加补考，但补考成绩只有合格和不合格两个等级。

3. 考试科目的满分均为 100 分，考查科目的满分均为 20 分。

（三）成绩处理和使用

1. 成绩处理

普通高中学生学业水平监测考试的成绩评定采用等级制，其中考试科目按照 A、B、C、D（分别为优秀、良好、合格和不合格）四个等级评定；考查项目按照合格、不合格两个等级评定。

2. 成绩使用

监测考试结果，对高中学校的教学具有指导和调控作用，是促进学校改进工作，提高质量的重要手段。监测结果是评估学校办学水平的重要依据，是评定示范性高中、重点高中的重要标准之一。监测结果也是各级教育行政、教研部门进行教育研究、指导和教学决策的重要依据之一。

学业水平监测考试成绩记入学生学籍档案（纸质与电子档案），并最终与每一位学生的综合素质评价报告并列呈现。

具有普通高中学籍的学生参加学业水平考试，每门科目成绩合格，且学分达到 144 分、综合素质评价合格，颁发普通高中毕业证书。不具有普通高中学籍的学生参加学业水平考试并成绩合格者，发学业水平考试合格证书。普通高中学业水平统一监测考试成绩可作为高考录取基本条件（具体办法详见高考方案），是推荐保送生必要依据。

（四）命题原则和命题范围

1. 命题原则

普通高中学业水平统一监测考试的命题以课程标准为依据，考查的内容和要求应与各科课程标准相一致。具体地说，学业水平统一监测考试应遵循以下原则：

（1）基础性。学业水平统一监测考试要注重对基础知识和基本技能的考查，杜绝超过课程标准要求的拔高。

（2）科学性。学业水平监测考试命题要做到试题设计与课程评价目标相一致，体现学业水平监测考试的功能和特点。

（3）创新性。试题设计突出新情境、新材料、新设计，注重考查学生综合运用基础知识的能力，力求联系学生生活经验和社会实际和关注科学技术的发展。

（4）公平性。实现试题素材和解答要求对于所有考生的公平性，避免偏题、怪题，认真考虑城市和农村的不同教学条件和能力，避免需要特殊背景知识和特殊解答方式的题目。

2. 命题范围

学业水平统一监测考试考查高中新课程规定的该科目必修学分规定的学习内容（各学科必修模块和毕业必修学分要求的选修模块）。

## 六、学业水平监测考试的改革构想

1. 高中学业水平测试的重点应该是改变过去过分关注学生学业成绩的做法，重视对学生多元能力和潜质的检测，促进学生全面、健康、和谐地发展。

学业水平测试侧重于“水平”考试，而非偏重于选拔和甄别。应立足于改革学生评价的目的、内容、主体和方法，以促进学生发展为目标，从德、智、体、美等方面综合评价学生的发展。不仅关注学生的学业成绩多少分，而且要发现和发展学生多方面的潜能和能力，了解学生发展中的需求，帮助学生认识自我，树立自信，充分发挥评价激励、诊断和促进学生全面发展的教育功能，形成学校、家庭和社会一体化的教育格局，促进学生综合素质的提高和身心健康发展。

2. 高中学业水平测试应当是落实普通高中课程方案，实施素质教育的重要保障。

通过学业水平测试反映学生在不同学习领域中表现出的知识

与技能、过程与方法、情感态度价值观的状况，创设有利于新课程实施的教育环境，推动普通高中课程改革，推动学校实施素质教育。

3. 高中学业水平测试应当为学生减负，而不应成为学生肩头的又一个考试重担。

尽管一直在提倡素质教育，然而现阶段我们国家的特殊国情，致使应试情结在短时间里不可能消失。在众多方案中，水平测试都安排在高二下学期时间，一口气考 10 门，接着就是准备高考，对大多数学生来说还是很有挑战性的，在形式上类似于“小高考”。因此如何让学生轻松地应对他们在高中三年学习成果的检验，还要做深入探讨。

# 第十二章　我国高考改革评述

高考是我国教育体系中的一个重要环节，是连接我国基础教育和高等教育的桥梁。自1977年恢复高考制度以来，我国基础教育在“高考指挥棒”的引领下取得了很大进展，为加速我国高级专门人才的培养做出了巨大贡献。但随着社会发展对高考功能需求的变化，我国传统的高考模式在影响素质教育健康发展方面，逐渐暴露出许多弊端。为适应教育和社会发展的需要，我国一直在不断地调整高考制度，各省市也积极探索高考改革措施，逐步涌现出诸如“3+1”、“3+2”、“3+X”，“3+1+X”等高考方案，这无不反映了我国教育的发展与前进。我国新一轮课程改革的实施给高考改革提出了新的要求，必须从高考的内容、方式、命题等多方面进行改革，才能适应新课程的发展要求。

## 第一节　我国高考制度的性质与功能

建立于1952年的高考，是高等学校招生考试发展到一定阶段的产物，是适应中国的政治、经济和文化、国情而建立的招生考试制度。它既反映了国家和社会的意志，也反映了高等学校的意志。建立50多年来，它不仅为我国国家建设培养了大量人才，同时对于维护社会公平、保持社会流动、协调社会关系发挥着重要作用。

## 一、高考的性质和特点

高考是典型的选拔性考试，又称常模参照性考试，它是“由合格的高中毕业生参加的选拔性考试。高等学校根据考生的成绩，按已确定的招生计划，德、智、体全面衡量，择优录取。因此，高考应具有较高的信度、效度，必要的区分度和适当的难度”①。

高考作为国家级的选拔性考试，是衔接基础教育和高等教育的桥梁，是教育体系中的一个重要的中间环节，对高等学校培养人才具有基础性作用，对基础教育具有导向性作用。它一方面要使高校的新生具有比较扎实的文理基础，并为高校不同学科和专业选择具有不同能力倾向的学生；另一方面，高考要能以较好的评价导向促进学生全面发展并产生积极的影响，给中学阶段的教育以更全面的导向作用。因此，与一般的考试相比，高考还具有以下特点：

（一）高考是区分考生差异的选拔性考试。与一般的目标参照考试不同，高考作为常模参照性考试，不是考核考生是否达到了高中教学的目标要求，而是满足选拔高等学校培养具有创新精神和实践能力的高素质专门人才和拔尖创新人才的需要。

（二）高考不是全面检查中学教学效果的考试，而是结合基础知识着重考核学习能力。心理学的研究和实践都表明，在已掌握的知识能与大学教学内容相衔接的条件下，一个人的智力特别是学习能力与大学阶段的学业成就相关程度最高，因此，高考应着重考核考生的学习能力。这一点在考试大纲中都有明确论述：“高考要以能力测试为主导，考查考生对所学相关课程基础知识、基本技能的掌握程度和综合运用所学知识分析、解决问题的能力”②。

---

① 中华人民共和国教育部．2005年普通高等学校招生全国统一考试大纲，2005，3.

② 教育部考试中心．2006年普通高等学校招生全国统一考试理科综合科考试大纲［M］．北京：高等教育出版社，2006.

（三）高考是大学教育工作的第一个环节，因而，高考并不是中学教育的一部分，考试内容也不等同于学习内容。高考是为大学招生而举行的考试，它的应试对象是已经从中学毕业了的青年，因此，它是大学教育工作的第一个环节，并不是中学教育工作的一部分。但因为高考是一种选拔性的考试，既然是选拔，就必然有竞争，竞争的结果就导致广大中学师生更多地围绕高考进行教与学，在无形中发挥着“指挥棒”的作用。

（四）高考的分数是相对分数，而不是绝对分数。高考的分数，其作用主要在于考生间的横向比较，借以区别考生的优劣差异，而不在于反映考生达到中学教学目标的程度，因此，它是一种相对的分数。

相对分数一般采取等级制，而高考却采取百分制，这是为了更精细地区分考生间的差异。由于高考评出的是相对分数，人们最关心的不是考生分数的绝对数值，而是每个考生的分数在考生全体中的相对位置，因此，高考最要紧的是把握好考生成绩的相对分布，而不是把握好平均分或合格线。

## 二、高考的功能

从表面上看，高考不过是高等学校招新生的考试，但由于其具有深刻的社会、经济、文化根源，承担着巨大的社会责任，往往牵一发而动全身。它作为一种分配教育机会和预演分配社会地位和社会资源的机制，实际上承担着教育筛选和社会筛选的双重功能。它除了为高校选拔新生、促进中小学生努力向学、规范中学办学方向以外，高考又是维护社会公平、保持社会流动、稳定社会秩序的重要手段，因而具有教育、社会和个体等多种功能。

（一）教育功能

高考是典型的选拔性考试，这种考试注重考试的区分度，以便从全体考生中选拔出特定数量的优胜者。作为一种竞争激烈的

大规模选拔性考试，高考的教育功能主要有以下三点：

第一，为高校选拔合格新生，保障高校生源质量。考试是一种测量工具，高考让所有应试者接受相同的挑战，将个人的才学和能力放在首位，因而历来被视为可以客观公正地选取优秀人才的公平尺度。与推荐制下新生良莠不齐、文化程度较低的情况大不相同，高考制度建立50多年来，特别是恢复高考30多年来，为高校选拔了大量合格的人才。他们中的许多人已成为现代化建设的中坚力量，“文革”后一度担心的文化断代和教育断代问题逐步得到解决。高考以其较强的信度和效度，对保障高校生源质量起到了重要的作用。

第二，促进中小学生努力向学，提高民族文化水平。高考作为选拔性考试，具有强大的以考促学的功能，考生水平全面提高后，为了保持一定的区分度，不得不再逐步提高试题难度，这种“水涨船高”的关系，促使高考试卷难度和考生水平节节上升。高考始终是激励考生奋发学习，高中不断改进教学的最实际、最强大的动力。许多参加过高考的过来人认为：“高考以它的公正，丰富了一代又一代人的梦想，赋予我们公正平等的竞争精神，冷静沉着的竞争意识，以及一种永恒的拼搏精神”。不少人虽经历过高考失败，但由于受过高考总复习的煎熬，养成了坚毅的个性，得到了一笔精神财富。

第三，引导、规范中学的办学方向和教学方向。高考与中学教学的关系理应是教什么就考什么，但高考对中学教学又有强大的制约作用，实际上往往是考什么就教什么，这就使高考具有指导中学办学和教学的功能。从某种意义上说，高考是国家意志和政治与学校教育活动之间的一个中介，通过高考的科目设置和命题设计，国家可以将政治理论和思想意志贯彻到中学教育中去。比如在政治科考试中包括时事内容，便可以引导考生关心时事。也可以通过科目增减和考试内容的比重调整，来调控或促进某些

学科的发展。高考指挥棒如果运用得当，便可以将学生引导到身心健康发展、注重能力和素质的方向上去。因此高考对实施教育方针、引导办学方向具有重要作用。

然而，高考是一把利弊兼具的双刃剑。高考在发挥正面功能的同时，因高考而造成的异化现象也十分显著，如中学只抓智育而片面追求升学率、文理偏科、学生负担过重等。

（二）社会功能

高考不仅是一种教育行为，而且是一种社会行为，一种社会大动员。如果说高考的教育功能是积极与消极错杂并存的话，那么其社会功能则以正面的居多。

第一，社会安定功能。维护社会公平、平等竞争是考试制度的灵魂和根本，“公平、公开和公正”是考试制度的核心理念。有学者认为，从法理的视角讲，考试制度其实可以视为一种崇尚“程序正义”的程序制度，正是基于“程序正义”理念的运作，赋予了古代科举制度和现代高考制度毋庸置疑的权威性。尽管现行高考制度受到一些人的质疑及现实的挑战，但有一个事实是无法否认的：高考已经成为当代中国最成熟和最权威的人才选拔机制，迄今尚没有任何一种制度可以取而代之。“文革”后高考制度的恢复并不止是使教育恢复了正常秩序，更是带动了整个国家由乱而治。而且，恢复高考是在粉碎“四人帮”这个大的政治背景下被作为拨乱反正的突破口而实行的重大举措，它使中国的人才培养重新走上健康的轨道，在使社会由乱而治的过程中发挥了至关重要的作用。由此可见，高考对恢复和稳定社会秩序具有重大的作用。

归因于高考制度的长期实行，在相当程度上形成了一种无论贵贱贫富，“在考试面前人人平等”或“在分数面前人人平等”的观念和社会文化氛围。实际上，千百年来，中国人都是不患寡而患不均，在考试方面则是不怨苦而怨不公。因为高考提供了公平

竞争的机会，在激烈的考场较量之后，名落孙山者虽然会有挫折感，但多数考生只能怨自己水平不够或运气不佳，无法怪他人或社会，即使是离分数线仅差一两分亦如此。然而，如果是因为不公平竞争而落榜，他们怨的则是政府和社会。在推荐制下，走后门盛行，无法上大学的学生眼见有权有势者的子弟能够接受高等教育，极易感到无助和愤怒，他们的不满就可能聚集起来针对整个社会，形成对现存秩序的反抗力量，影响社会的稳定。在取消报考年龄限制以后，高考理论上成为一种终身考试，它为每一个落第者始终保留着下一次成功的机会与希望，这就为落榜考生提供了更多的选择，也使一些考生和家长的不满和失望不至于积累到危险的程度，因此高考在一定程度上起到了社会的“减压阀”与“稳定器”的作用。

第二，社会流动功能。世界上各国的大规模教育考试都有强大的评价和筛选功能，也各有其重要的作用和影响，在中国表现得尤为突出。

社会流动分为社会阶层流动和区域流动。考试能在相当程度和范围内促进社会阶层流动，这在科举时代已得到证明。传统社会所谓“茅屋出公卿”，“朝为田舍郎，暮登天子堂”等格言，不仅是统治者对读书人的利诱和鞭策，也是当时科举造成社会阶层流动的历史的真实写照。科举时代的社会阶层流动能使社会保持一定的活力，使官僚结构不断得到更新。在现代社会，社会流动与政治稳定之间也存在着一定的关系。许多政治分析家认为，一个工业社会的政治稳定或不稳定由其社会流动率所决定，高流动率的社会是稳定的，反之则是不稳定的。高考在促进社会阶层流动方面起着重要的作用，参加高考长期以来是中国农村学生获得城镇户口的一条重要渠道。虽然在城乡差距较大的情况下，高考在促进农村人口向非农业人口转化过程中，会使本来人才就稀少的农村人才更少，但它客观上促进了农村教育的发展，成为激发

农村人口接受教育的无形动力，也在一定程度上促进了农村城市化进程的发展。在中国高等学校尚未全面收费之前，高考促进社会阶层流动的作用尤其明显，多年来一直是许多农村青年改变命运、进入社会上层的主要甚至是惟一办法。托福于高考制度，许多家境清寒的学子经过大学深造，现今已成为中国各个部门、各个行业的骨干力量，这是高考的社会流动功能的生动体现。

社会水平流动是指区域流动。区域流动（或区位流动）与阶层流动相关，区域流动可以说是社会阶层和职业流动的必要条件，也可以说是它们的先决条件。居住地区的改变往往是一个人生活和事业进程中的组成部分。由于高考长期实行全国统一考试，考生在当地就可以方便地报考全国重点大学或跨省区录取的大学。这一制度造成大范围的人才流动，为这些高校选拔了来自于不同省市、不同民族的人才，大大开阔了学生的视野，有利于各地文化的交流和国家的统一。比起过去高校单独招考时一般家境不好的学生较难到大城市报考著名大学的情况，高考制度在促进人才区域流动方面起着重要的作用。有研究者以云南一个省、一个少数民族自治州、一个县的实例和全国高校招生录取年度统计分析，证明高考是人才从农村向城镇流动或区域间流动的重要渠道，不仅促进了边疆地区、民族地区的教育发展，而且比较公平地控制、调节着向上流动的通道，深刻地影响着中国的社会结构。由此看来，高考的功能相当复杂，在发挥其教育功能的同时，也发挥出巨大的社会功能，我们在进行高考改革时，就应顾及高考改革的社会影响。

（三）高考的个体功能

高考是一种社会性的制度安排，这种制度的实施必然对受教育者个体产生影响。高考内容和高考方式甚至可能引起某些个体命运的起伏。具体而言，从个体层面看，高考具有以下功能。

1. 分配个体生存资源，促进个体发展的功能

从哲学意义上说，人既是历史和社会的结果，又是历史和社会的原因；既是社会的客体，又是社会的主体，二者统一于人的实践活动中。

从文化意义上说，个体作为文化人，有必要在文化能力、文化产品和文化体制方面获得自己的必要位置。高考，直接指向于教育机会的分配，成功者必然提高自己的文化能力，而成为文化的活载体，并享有和创造文化产品，从而在主流文化中占据有利位置。高考分配高等教育机会，这是不证自明的问题，但在更深层的意义上，它还具有其它功能。高考除了分配教育机会、预演分配社会机会、社会身份地位外，还再生产社会结构和形成新的社会资本或文化资本。

正由于在个人的知识资本、社会资本或文化资本形成中的重要作用，高考实际上在发挥着分配个体生存资源方面的功能。

2. 提高个体享用能力的功能

高考实际上在分配高等教育机会和社会地位。高考的成功者，必然在资源利用和社会地位的占有上成为优势群体。这种优势，不仅仅使他们获得了更好的教育机会，具有更好的发展基础和拥有更大的发展机会，而且在社会心理上可以提高他们社会地位的合理性和合法性。

高考作为国家举办的选拔性考试，它在积极发挥社会、政治和个体功能的同时，也因其所扮演的特殊角色，备受世人瞩目。但也正因为此，对高考的性质和功能认识也愈应该重视。

首先，高考为政治服务，但不等同于为政治运动服务；其次，甄选不等于评价。高考是对考生进行德、智、体全面甄选，为高等学校挑选新生实施的社会评价活动。它的设计、实施和评价等活动都应指向并服务于高校选拔新生这一基本目标，选拔标准只能是考生是否具备接受高等教育的基本条件而不直接对中学办学

水平以及学生其他素质、能力和水平进行评价。第三，高考竞争不等于应试教育。有考试就有竞争，有竞争才有活力。这是高考的功能和性质决定的。但是考试并不必然导致应试教育。第四，高考对人才培养具有较强的调控功能，但不能将高考当作万能的调控工具。

## 第二节 我国高考改革的历史回顾

1949 年新中国成立不久，就确立了高等教育全国统一招生考试的制度。从此，高考这个词就一直伴随着国人，牵动着国人的视线和神经。我国的高考历史，可以说是一部跌宕起伏、艰难曲折的历史。它伴随着我国政治、经济和高等教育的发展，走过了在改造中建立、在建立中改革、在改革中发展的道路。经过 50 多年的发展，目前我们已初步建立起具有中国特色的社会主义高考制度。纵观高考的发展历程，大致可分为以下 6 个不同的发展阶段。

### 一、建国初期过渡阶段（1949～1951 年）

此时的教育是以改造旧教育体系，建立新教育制度为主要特征的过渡时期的教育，高校招生考试制度也处于过渡阶段。

1949 年 9 月份《中国人民政治协商会议共同纲领》提出“有计划有步骤地改革旧的教育制度、教学内容和教学方法”的指导精神，新生的人民政府提出“维持现状，立即开学”的方针。除北京大学、清华大学、南开大学、北平师范大学及北洋大学等少数几所高校在 1949 年实行非实质性的联合招生外，全国其他高等院校都实行单独招生。招生计划、招生条件及招生办法都由各校自行决定，各校享受了最大限度的招生自主权。但整体的录取率都比较低。

1950 年 5 月 26 日教育部发布《关于高等学校 1950 年度暑期招考新生的规定》指出，本年度高等学校招生，由各大行政区分别在适当地点定期实行全部或局部的联合或统一招生，并允许各校自行招生，各校招生名额由各大行政区负责审核。各校招生不足额的情况得到了很大改善。

1951 年教育部发布了《关于高等学校 1951 年度暑期招考新生的规定》，其中沿用了 1950 年招生规定的基本内容，但对统一招生范围和程度做了新的补充。即“为进一步改正各校自行招生所产生的混乱状态，各大行政区分别在适当地点争取实行全部或局部高等学校统一或联合招生，全国统一考试日期；如有困难，仍允许各校单独招生；在其他地区招生时应尽量采取委托的方法进行”。当年统一招生规模迅速扩大，全国 214 所高校中参加统一招考的达 149 所，招考时间也更为集中，基本达到了“全国统一考试时间”预期目标。

## 二、统考制度建立阶段（1952～1965 年）

这一阶段我国正处于社会主义改造基本完成、社会主义建设全面展开的时期。根据政治和经济形势，尤其是高校的院系调整形势，教育部对 1952 年的高考作出了新的规定：在中央一级成立全国高等学校招生委员会；全国所有高等学校除个别经教育部批准者外，一律参加统一招生，采取统一领导与分省、市、自治区相结合的招生办法，招生名额应报请审核批准，严格禁止乱招乱拉，招生日期和考试科目全国统一规定。至此，统一的高考制度正式建立。

1954 年 5 月 28 日高教部、教育部联合发布招生规定指出：今年招生必须贯彻中央统一计划、大区组织执行，并由各校直接负责审查录取新生的组织方针。录取实行“按总分由高分到低分，分为若干等级，逐级分配”的办法。

1957 年 4 月 24 日高教部、教育部发布招考规定，要求在高校招生应进一步照顾考生的志愿，录取新生一般不采取计划分配的办法，按考生志愿不能录取满额的学校，再采取时要向考生征求志愿或者采取再次报考的办法来补足缺额。

1958 年 4 月 4 日，中共中央发出《关于高等学校和中等枝术学校下放问题的意见》。提出改变统一招生的制度和毕业生分配办法。同年 7 月 11 日，教育部发布年度高校招生规定，进一步明确“改变全国统一招生制度，实行学校单独招生或者联合招生”，1959 年又发文“改变 1958 年各校单独招生或联合招生的办法，恢复全国统一命题、一次考试、分批录取的办法”。这一高考体制一直延续到 1965 年。

**三、废除高考统考阶段（1966～1976 年）**

1966 年，中共中央国务院发出《关于改革高等学校招生考试办法的通知》，指出现行的招生办法必须彻底改革。并且决定从当年起高校招生工作下放到省、市、自治区办理；高校招生取消考试，采取推荐与选拔相结合的办法。1970 年大部分高校恢复了招生工作，但取消了文化考试，而采取“自愿报名、群众推荐、领导批准、学校复审”相结合的招生办法。由于文革的原因，致使高等学校连续 4 年停止招生，废止全国统考达 11 年之久，造成一个时期各种专门人才匮乏，严重影响了我国的经济建设和社会发展。

**四、恢复高考阶段（1977～1984 年）**

经过十年文化大革命的浩劫，中国的教育界百废待兴，迫切需要整顿和恢复。1977 年 6 月教育部在山西太原召开了“一九七七年高等学校招生工作座谈会”；同年 8 月，邓小平在科学和教育工作会议上说：“今年就要下决心恢复从高中毕业生中直接招考学

生，不要再搞群众推荐”；10 月 12 日，国务院批准了教育部《关于1977 年高等学校招生工作的意见》，正式恢复高等学校招生统一考试制度。

1977 年 11～12 月，各地高考陆续进行，考试由各省、自治区、直辖市命题。当年共有 570 万人报考，共录取 27 万人，录取率4.7%（27:1）。1978 年7 月 20～23 日，全国高等学校开始实行统一招生考试。1979 年7 月 7～9 日，全国高等学校招生统一考试举行。当年共有 468 万人报考，共录取 28 万人，此后高考就定为每年的7 月 7～9 日举行，报考及招生人数逐年上升，我国高考招生制度逐渐走上正轨。

### 五、改革起步阶段（1985～1993 年）

1985 年 5 月 27 日，中共中央发布《关于教育体制改革的决定》，规定“改革高等学校的招生计划和毕业分配制度，扩大高等学校办学自主权”，实行国家计划招生、用人单位委托招生、招收少数自费生3 种招生办法。同时，随着对外开放的不断扩大，国外教育测量学理论等一些考试方法进入我国，冲击着我国传统的考试手段。

至今人们耳熟能详的有：1984 年特殊专业（艺术类等）招生实行统考 + 单考，实行保送生制度；1985 年，开始试行高考标准化改革、会考制度；1991 年“三南”（湖南、云南和海南）高考科目改革……中国高校招生制度开始进入了一个新阶段。

### 六、深化改革阶段（1994 年至今）

1994 年以来，高考有两项比较大的制度性改革。一是将多种招生计划形式合并为一种计划形式，通常称为招生计划并轨。这是我国招生分配制度的重大改革，它于1994 年进行试点，1997 年在所有院校全部实行。二是自 1999 年启动的新一轮高考改革。

1999年2月，教育部出台了《关于进一步深化普通高等学校招生考试制度改革的意见》，其指导思想是：在改革中要始终坚持有助于高等学校选拔人才、有助于中学实施素质教育、有助于高等学校扩大办学自主权的三项原则，把高校招生考试工作提高到一个新水平。从此掀起了中国高考的新一轮改革。

改革的内容主要有四个方面：一是高考科目设置改革，计划用三年左右的时间推行“3+X”科目设置方案。二是高考内容的改革，这是改革的重点和难点。命题要以能力立意，注意考察跨学科的综合能力。三是高考形式的改革，积极探索一年两次考试的方案。四是录取方式的改革，重点是实施计算机网上录取。

回首这段改革历程，至今历历在目。如1994年，在会考基础上高考开始实行“3+2”科目改革；1999年规定保送生都必须参加综合能力测试；2000年试行春季高考；2004年试行高考分省命题；2007年四省区实行新课程高考方案等。

## 第三节　我国高考改革的方向和目标

2004年教育部下发《2003－2007年教育振兴行动计划》文件，为我国当前高考改革指明了方向和目标，文中明确提出“结合新课程的全面推进，深化高考内容改革；推进高考制度改革，进一步建立以统一考试为主、多元化考试和多样化选拔录取相结合，学校自我约束、政府宏观指导、社会有效监督的高等学校招生制度。完善高等学校招生网上远程录取系统和网上阅卷系统，建设招生信息化管理与服务平台”。

### 一、深化考试内容改革

命题要把以知识立意转变为以能力立意，转变传统封闭的学科观念，在考查学科能力的同时，注意考查跨学科的综合能力。

内容要更加贴近实际与生活，淡化解题技巧，更加注重应用能力考查，同时寓思想道德教育于考试之中，引导积极的价值导向，在一定程度上体现地方特色。这一趋势还将继续进行。

## 二、继续稳步推进分省命题的改革

多年来，“统一命题、统一考试”的高考模式，因其具有公平、高效、权威的特点，成为社会最为认可的高校新生选拔方式。但这种简单化的“分数面前人人平等”的考试理念与考试方式，因为无法兼顾各省经济、文化与教育水平的差异，在高等教育大众化与新课程全面推行的背景下，面临着许多压力与挑战。在这种背景下，从2000年开始，教育部首先在上海、北京等城市进行了自主命题的试点，到目前为止，全国实施自主命题的省份达到16个，考生人数约占全国考生总数的2/3左右，“统一考试、分省命题”的格局已经基本形成。

分省命题因为充分考虑到各省的经济、文化教育的发展水平和课程改革的实际，是在高考改革坚持统一的基础上，走向多样化的关键步骤之一。它不仅为新课程改革、推进素质教育提供了相对宽松的外部评价机制，也为高考正确发挥“指挥棒”的作用提供了空间。同时，由于我国大学新生录取实行“分省录取”的模式，考生的竞争事实上主要来自省内。因此，采取“分省命题”的方式在与录取模式相一致的基础上，也在一定程度上降低了考生的竞争压力和心理负担。

更为重要的是，分省命题作为考试权力下放的重要形式，为考试管理提供了新的思路与途径。在某种程度上，它为各省逐步建立起成熟的考试命题、考试管理队伍提供了契机，为进一步推进考试的改革乃至非营利性考试专业机构的建立埋下了伏笔，为促进考试机构之间的良性竞争、推行“服务考生”的理念打下了基础。可以说，分省命题是涉及高校招生考试乃至教育全局的重

大制度性改革。此外，分省命题在客观上也降低了高考试题泄漏的风险。在全国统一命题的模式下，一旦发生试题泄密的情况，影响范围将波及到全国的考生，而在分省命题的背景下，其影响范围则缩小很多。

### 三、不断完善多元化的选拔录取考试

从多年的高考实践来看，统考是高考选拔人才的基本主体形式，其科学性、公平性和效率性是不容置疑的。不过，在教育日益以人为本的今天，也要求考试能对考生综合素质、个人特长进行选拔。这方面，在统考之外已有多项政策、多种形式发挥了补充作用，从保送生、特长生等直至自主选拔招生，都为不拘一格选人才提供了空间。因此，多元化的选拔录取考试将稳中求进、不断探索、不断完善。

### 四、积极探索多层次考录改革

在高等教育迈向大众化的今天，庞大的高教系统中不同层次的学校有不同的办学目标和定位以及在人才培养上的层次性，这决定了选拔性考试的分层次性。现在，一张考卷既考进北大、清华的科技精英，又考接受职业教育的高技能劳动者，而随着高等教育大众化的加快，分层次考录改革的迫切性也越发突出。因此，改革将通过从形式或科目的不同来积极探索和实践对本科和高职的多层次考录改革。

### 五、高考手段现代化

随着信息技术的不断普及，基于计算机为主的信息技术也开始渗透到高考的各个环节当中，高考阅卷、考生档案以及考生录取等相关环节都可以通过网络计算机技术进行。目前我国基本实现了高等学校招生网上远程录取和网上阅卷，建设招生信息化管

理与服务平台基本成型。

为解决人工录取存在的劳动强度大、速度慢、差错率高等情况，1999年教育部在全国组织了10个省（区、市）和北京大学等400余所高校进行了网上录取试点工作。目前，高考网上录取基本普及。实施计算机网上录取新生，是普通高校招生的一项重要改革，可以说是录取方式的一次革命。计算机远程网上录取的实现，不仅更大程度地发挥了网络优势，提高了工作效率，而且净化了录取环境，使录取结果更加公平、公正；同时，实行计算机网上录取，顺应了现代信息技术发展的趋势，通过实行计算机网上录取，高校取得学生的基础数据，有利于高校建立招生、学籍管理、毕业就业服务一体化的学生管理系统，从而推进高校学生管理的现代化进程。今后，计算机在高考中的应用将更加广泛。

### 六、综合素质评价逐渐被纳入高考范畴

综合素质评定和综合实践活动课程是新课程方案的一个亮点，在当前中考改革中，初中课改实验区已积极将初中学生综合素质评定结果引入中考，取得了良好效果。随着高中新课程方案的不断深入，这一特色也将逐步包含在高考评价范畴。在2007年首批新课程改革高考试点省市中，除山东外，广东、海南、宁夏三省（区）都在新方案中对综合素质提出了明确要求。

如广东省方案规定，综合素质评价由省教育厅公布统一的评价要求，由学校实施。综合素质评价内容以写实性文字表述方式分项目记入考生电子档案。内容包括中学期间各科目学分获得情况，技术领域、艺术领域、体育与健康领域、综合实践活动的表现情况或获奖情况等，以及取代现行考生电子档案的毕业考试成绩、毕业鉴定，录取时供高等学校参考，作为高校录取依据之一。

可以预见，随着课程改革和高考改革的不断深入，单纯以高考学科分数为录取标准的高考形式将会得到不断改变，学生的平

时表现、手工制作、情感态度等体现综合素质的内容也将逐步被纳入到高考评价的范畴当中。

### 七、高考形式的多样化

为改变"一考定终生"的不合理之处，教育部从2000年1月开始，就在北京、安徽、上海等省市试行春、夏季高考，积极探索一年两次考试的方案；同时，部分省市已经在中考中实行二次考试形式，这对我国高考形式多样化提供了一定的参考依据。

## 第四节　新课程背景下我国高考的命题特点与趋势

高考作为一个全国性的选拔考试，既要担负起高校选拔人才的重任，又要为中学教学、新课程实施提供重要的导向作用。纵观首批新课程实施省市的高考命题，可以从中发现以下规律：

### 一、合理控制试卷难度

高考试题在与中学实际教学的关系上较往年密切，能够较好地反映我国目前大多数中学的教学实际状况。这表明我国目前的高考改革正逐步探索出适应于中小学课程改革的命题方式，为今后高考改革的顺利进行奠定了良好的基础。

### 二、进一步加大试题的跨度

主要表现为试题切入点较低，层层深入，为考生搭建台阶明显；难点分散并且在难度安排上进行人性化处理，由易到难安排；试题难度跨度进一步加大并出现一些社会化试题，体现出考生水平。

### 三、客观题减少，主观题增加

题目总量减少，增加考生思考时间，加大思维考查力度。

**四、试题回归基础，回归教材**

重点考查学科的主干知识，从而淡化学科技巧和常规难点。

**五、增设创新性试题，落实新课程理念**

创新题强调综合性和深层理解；增设开放性的试题；提出新的设问角度；把考点设置在考生容易出现的错误上；注重与地方特色相结合。

由此可见，高考命题的趋势可以归纳为将具体知识能力化，将学科问题综合化，将实际问题学科化，将大学内容初等化；考核内容本质前沿，试卷结构稳定开放，命题手段科学创新，答题要求重想轻算。

## 第五节　我国首批高中新课程实验区高考改革方案简介

2004年秋季开始，山东、广东、海南、宁夏四省区被确定为全国首批普通高中新课程的实验省市。作为首批进入新课程高考的这四个省区，新课程背景下的高考怎么考，成为社会普遍关注的热点。结果表明，这四个省区都交出了较为满意的答卷。

**一、山东：3+X+1**

2007年山东省采用“3+X+1”新课程高考方案。山东省是我国高中新课改实验区最先出台高考方案的省市，因而其高考方案备受人们关注。

这里的“3”指的是语文、数学（分文/理）和外语三个科目，分值各为150分，是所有考生必考科目；“X”代表文科综合或理科综合，分值各为240分。文科综合包括思想政治、历史、地理三个科目的必修内容和部分选修内容，理科综合包括物理、化学、

生物三个科目的必修内容和部分选修内容。报考文史类、艺术文科类考生须参加文科综合的考试，报考理工农医类、体育类考生须参加理科综合的考试。“1”指“基本能力”测试，包括技术、体育、艺术、综合实践等，以及运用所学知识解决生活和社会实际问题的能力，卷面分值为100分。“1”的考试内容不分文理科，是所有考生必考科目。

试卷结构方面，增加了选做题，各科试卷分卷Ⅰ和卷Ⅱ，卷Ⅰ是必做题，以选择题题型出现，主要考查必修内容；卷Ⅱ含必做和选做题，是以非选择题形式出现，必做题考查必修内容，选做题考查选修内容。

基本能力测试的必做题主要涉及技术、艺术、体育与健康、综合实践活动4个学习领域的必修内容，选做题分人文与社会、科学两个部分，内容分别涉及这两个领域的必修内容。为防止偏科现象，方案规定：文史方向的考生只做科学部分的题目，理工方向的考生只做人文与社会部分的题目。

在全国尚无经验可借鉴的情况下，山东省高考方案的出台无疑是一个大胆的创新和突破。应该说本方案较好地解决了一个省率先改革与大环境的接轨问题，充分体现了“减轻学生负担、纠正文理偏科、适当加强外语、引导学生个性发展、突出学生素质和能力考核”等改革意图。山东省新课程高考方案的出台，也为其他新课程试点省市高考方案的出台提供了重要的参考。

## 二、宁夏：3+小综合

2007年宁夏回族自治区高考采用“3+小综合”模式，并将学生综合素质评价作为高等学校录取参考。

“3”指语文、数学和外语三个科目，是所有考生必考的科目。“小综合”指文科综合、理科综合。文科综合包括政治、历史、地理三个科目，文史方向的考生必须参加文科综合考试；理科综合

包括物理、化学、生物三个科目，分值300，时间150分钟，理工方向的考生必须参加理科综合考试。综合素质评价：中学对考生的综合素质做出评价，评价指标需包括技术、体育与健康、艺术、综合实践等四个学习领域。综合素质评价以文字形式呈现，通过电子化的方式提供给高等学校作为录取参考。

试卷分Ⅰ、Ⅱ卷，考试内容分必考和选考，其中第Ⅰ卷为三个科目的必考题目，第Ⅱ卷由必考和选考题构成。试卷按题型、内容和难度进行排列，选择题在前，非选择题在后，同一题型在同一学科的试题相对集中，同一学科中不同试题尽量按由易到难的顺序排列。

### 三、海南：3+3+基础会考

2007年海南省高考新方案实行“3+3+基础会考”模式，具体如下：

“3+3”科目组合是指：文史类的语文、数学（文）、外语（加听力）、政治、历史、地理；理工类的语文、数学（理）、外语（加听力）、物理、化学、生物。

基础会考是高中毕业水平考试。其中文科类的基础会考科目包括物理、化学、生物、通用技术和信息技术；理工类的基础会考科目为政治、历史、地理、通用技术和信息技术。依据课程标准，文科类与理科类的数学科目考试范围不同。体育类考试科目组合为：理工类考试科目组合，再加考相关学科。艺术类考试科目组合为：文史类考试科目组合，再加考相关学科。

“3+3”科目的考试成绩按卷面实际获得成绩计入录取成绩总分，基础会考成绩按卷面实际获得成绩的10%计入录取成绩总分。2007年高考总分比改革前增加了40分，达到790分。基础会考成绩，一次考试两年有效。2007年非应届毕业生的考生可参加当年举行的基础会考，会考成绩按规定比例计入考生当年录取成绩

总分。

**四、广东：3 + 文科基础/理科基础 + X**

2007 年广东为适应新课程高考要求，实行了“3 + 文科基础/理科基础 + X”的高考新方案。

“3”指语文、数学（分文、理）、英语。文科基础包括：思想政治、历史、地理、生物等学科课程标准的必修课程（部分）内容；物理学科课程标准的必修模块物理 1、物理 2 和选修 1－1 模块；化学学科课程标准必修课程的化学 1、化学 2 和选修课程的化学与生活模块。其中思想政治、历史、地理的内容占分值 70%，物理、化学、生物的内容占分值 30%，全部采用选择题。理科基础内容包括：思想政治、历史、地理、生物等学科课程标准的必修课程（部分）内容；物理学科课程标准的共同必修模块物理 1、物理 2 和选修 2－1 模块或选修 3－1 模块；化学学科课程标准必修课程的化学 1、化学 2 和选修课程的化学与生活模块或化学与技术模块。其中物理、化学、生物的内容占分值 70%，思想政治、历史、地理的内容占分值 30%，全部采用选择题。

“X”为专业选考科目，从物理、化学、生物、政治、历史、地理、音乐、美术、体育等 9 门学科中任选一科。专业选考科目考试内容为各学科课程标准的必修课程（部分）内容和广东省教育厅指定的选修课程（部分）内容。

## 第六节　我国高校招生制度改革探讨

高考是牵动千家万户的大事，它的恢复为国家合理、有效地选拔人才，为学生提供相对公平、公正的竞争环境创造了良好的条件，发挥了积极的作用。但随着高等教育大众化阶段的到来，现行高校招生制度的内在矛盾逐渐显现，高校招生制度的改革已

成为社会发展的必然选择。

## 一、高校招生制度改革的必要性

### （一）现行高校招生制度的弊端

改革开放以来，我国高校招生制度虽一直处于变革中，但弊端仍非常明显。

1. 高考评价方式存在局限性

高考选拔学生应该是多方面的，但现行招生制度仅以文化知识作为唯一的评价内容，难以全面衡量一个人的素质。

2. 高考“指挥棒”的负面效应

高考“指挥棒”的巨大牵动效应，导致基础教育向应试教育方向畸形发展。

现行招生制度下高校的招生自主权很小，学校的独立法人地位得不到体现；考生选择的权力未受到尊重，专业选择与就业间存在矛盾；学生在高中阶段的超负荷运行不适应高校的管理要求，为学生接受高等教育留下了后遗症。

3. 人为制造了区域间的不平等性

现行的高校招生是以省为单元组织的，招生政策和录取线标准由各省制定，不同省份的录取标准差别较大，发达地区与不发达地区的录取标准成倒差。

### （二）高等教育大众化的挑战

国际上通常把适龄人口中15%以下接受高等教育划为精英教育，15% -50%为大众化教育，50%以上属普及化教育。高等教育的大众化、普及化，已成为20世纪60年代以来世界高等教育发展的大趋势。1999年开始的高校扩招至如今，我国高等教育入学率迅速提高，进入高等教育大众化阶段。

1. 高等教育大众化需要多样化的招生制度

高等教育大众化的一个重要特征是多样化，包括学校类型多

样化、入学标准多样化、人才培养模式多样化、毕业生就业形式多样化。多样化的高等教育需要与之相适应的多样化的招生制度。传统精英教育认为高等教育是研究高深学问，只有要学习高深学术课程的学生才能接受高等教育。随着高等教育大众化时代的到来，高等教育在培养少数“学术精英”的同时，更要培养大量的应用型人才。注重“学术取向”的高考作为高校入学的最主要依据，已不能适应高等教育大众化的要求。改变单一的、以现行高考为主要依据的招生制度，建立一套适合不同类型、不同层次、不同培养模式的高校选拔人才的招生制度，已成为高等教育大众化阶段高校招生制度改革的目标。

2. 素质教育需要与之相适应的高校招生制度

高等教育走向大众化，改变了“千军万马过独木桥”的状况，为中等教育实施素质教育创造了条件。现行高考制度的改革，旨在革除旧弊、创新制度，更好地为社会经济发展服务。现行招生制度如不作根本性改革，“应试教育”仍难退出历史舞台。虽然高校招生规模的扩大，拓宽了高等教育的“入口”，但高等教育的“卖方市场”仍在相当长时间内存在，特别是高校的层次和专业的不同，依然会对中等教育产生影响。科学、合理的高校招生制度是中等教育实施素质教育的保证。

## 二、高校招生制度改革的原则

为适应高等教育大众化的需要，高校招生制度的改革必须坚持“三个有利于”：

（一）有利于高校科学选才

高校具有教学、科研、服务三大功能，具体体现在传递知识、发展科学文化、为社会服务三方面。高校为更好地发挥传递知识的功能，就要求学生具备获取知识的基础知识，要求高考能充分考察考生的知识以及运用知识的能力；高校为实现发展科学文化

的功能，要求高考内容能激发考生的学术思想和兴趣，能够检测考生的学术倾向和道德修养；高校为了服务社会，要求高考朝着有助于引导学生培养和开发知识转换能力、想象力和创造力的方向发展。选拔可育之人而培育之，是高校办学之目的。科学选才是实现这一目的基本保证。

（二）有利于高校自主办学

招生、培养、就业是高校工作的三大环节。扩大招生权限是扩大高校办学自主权的重要内容。随着高校逐步成为相对独立的市场主体，高校会选择利于实现高校功能和市场需要的学生，并提出相应要求。

（三）有利于中等教育实施素质教育

素质教育以人为本，从人自身发展的角度来规范教育。其基本思想是以人的身心发展为目的，注重学生的全面发展。因此，改革高校招生考试，使中等教育由应试教育走向素质教育，应是高校招生制度改革的出发点。

## 三、高校招生制度改革的具体设想

高校招生制度包括考试、录取、管理等几方面，是一个彼此紧密联系、相互制约、相辅相成的有机整体。高校招生制度的改革要遵循党的教育方针，坚持教育创新和素质教育的方向，促进人的全面发展。既要全面规划、系统推进，又要确立重点、分步实施，力争在几个重要方面或环节上取得进展。

（一）扩大高校招生自主权

扩大高校招生自主权，有利于高校根据自身实际情况以及考生所报专业确定招生人数，有利于学校招收有特殊才能的学生。根据目前我国高等教育改革的实际情况，可有选择地试行以下几种模式：

1. 考生参加全国统一高考，高校采用多元化的“评价录取”

模式。高校要探索并建立有助于反映学生综合素质和个性特长的多元化考试评价系统。可考虑将高考成绩作为评价录取的重要依据，同时兼顾考生的特长。如高考中某科成绩特别优异、能力证书、成果发明证书等；高中会考成绩；特殊专业的加试成绩、面试成绩等。

2. 高校自行制定录取标准和招生办法。根据各自的办学思想、专业特点，在多元录取的基础上，自行制定录取标准。可考虑在部分自律意识强、制度完善、管理严格的高校中，按照标准刚性化、程序规范化、办法公开化、结果公示化的要求制定招生办法，在招生计划一定数量范围内，在高考成绩达到批次投档线上进行综合评价、自主选拔的试点。

3. 为增加考生的专业选择机会，考生可同时报考若干所高校。通过招生咨询，考生对所选专业与职业的关系、自身个性与专业需求的协调程度做到胸中有数，并按个人意愿选择多所高校，选择符合本人特长的专业就读。高校对招生计划应具有一定的调配权，可根据学校专业和学生意愿进行适当调整。

（二）改革高考制度

1. 高考科目和考试内容改革

在推行“3 + X”高考科目改革时，“X”是多样性、选择性的，它代表“3”以外的其它科目，由高等学校自行确定。“X”代表了学生的选择权和高校的选择权，有利于高校自主确定招生人数，招收有特殊才能的学生。但在试行时，大多数省市只实行“3 + 文综/理综”，“X”已成了综合考试的代名词，没有很好地体现设置的初衷。因此，“X”必须转向小科目，由高校根据专业自主确定其中的某一科目，使人才选拔有较强针对性。另外，考试内容应更加注重对考生能力和素质的考查，试题内容要加强理论与实践的联系，试题设计增加应用性和能力型题目。

2. 高考次数改革

第二次全国教育工作会议要求高考每年举行两次甚至多次，这有利于克服“一考定终身”的弊病。从世界高校招生制度的发展趋势来看，两次考试制度比“一考定终身”具有更大的优越性。目前我国个别地区试行的两次高考制度并没有改变“一考定终身”的实质，春季高考只是形式，并未受到真正的重视，高中会考制度很大程度上也仅是“高考”模拟考试中的一次而已。高等学校招生制度如何从“一次考试”向“两次或多次考试”制度转变已是一个急需解决的课题。

（三）招生计划管理改革

招生计划管理应打破区域界线，实行国家和地方两级统筹管理。国家可把高校招生计划分为两块，一块为国家招生计划，另一块为地方招生计划。国家计划招生由国家根据全国招生总数、报考人数及考试情况，统一划定国家录取分数线（即国家线）；地方计划招生由各省根据本地区招生计划、报考人数及考试情况，确定地区录取线（即地方线）。招生批次统一划分，重点本科和普通本科批次在国家计划中完成；高收费本科与高专高职等批次在地方计划中完成。考虑到高等教育发展的区域不平衡性，部属院校的招生可适当向西部地区倾斜。

（四）加强高校招生管理

高校招生录取的自主权扩大后，高校录取管理工作就显得尤其重要。依法行政、依法治招，是深化高校招生管理制度改革的核心。要建立和完善各项规章制度，进一步明确政府、高校和考生三者之间的权利与责任，逐步形成高校自主招生、自我约束、地方招办监督服务，国家宏观管理的管理体制。政府要加快修订《普通高等学校招生工作暂行条例》，制定高校考试组织及处罚条例，变指令性管理为指导性管理。高校要规范并完善招生章程，严格履行招生章程的各项承诺，使招生章程成为高校规范招生、社会依法监督的重要依据。

# 第十三章　中学竞赛考试概述

竞赛，是通过制定一定的规则，以比赛的方式，使少数优秀者获胜的活动评价方法。竞赛活动适用于各个专业技能和各个学科学业的评价。其中，以竞赛的形式进行的考试称为竞赛考试。在学校教学中，举办的竞赛考试活动很多，比如智力竞赛、专题知识竞赛（如法律知识竞赛）、专项知识比赛（如作文比赛、编程比赛）等等。中学里常常举办的竞赛考试有作文竞赛、奥林匹克科学竞赛、英语竞赛、政史知识竞赛等。在目前我国的中学阶段，竞赛考试活动中涉及范围最广、影响最大的是奥林匹克科学竞赛。本章即主要以中学奥林匹克竞赛为例，介绍中学竞赛考试的相关知识。

## 第一节　中学竞赛考试的历史沿革

竞赛考试并不是时髦的现代产物，早在我国古代及中世纪的欧洲就出现了竞赛考试活动。随着竞赛考试的发展，中学竞赛考试得以在我国流行开来并广泛实施，对繁荣中学教育、促进学生能力的提高发挥了积极作用。

### 一、古代竞赛考试的雏形

伴随着考试制度的产生和发展，竞赛这种考试形式也早已在我国古代的考试制度中出现。据载，我国古代考试制度中很多都

是采用竞赛的模式，如何休《春秋公羊传解诂》注曰："行同而能耦，别之以射，然后爵之。"意思是说，若贤能在道德品行和技能相同的情况下，则以考核射箭技术高低来区别，授予优胜者一定爵禄。自隋唐时期开始延续了一千三百多年的科举制度，开考之年又称为"大比"之年，即所有考生"同台竞技"，特别是状元、榜眼、探花的产生更是经过层层比拼决出的少数优胜者。

十六世纪初，意大利的数学家们喜欢提出问题向其他数学家挑战，其中解三次方程那场比赛使人们记忆犹新。意大利数学家丰坦那（Niccolo Fontana），人称"塔塔里亚"（Tartaglia 意为口吃者），出身贫寒，自学成才，后以教书为生。1535 年意大利数学家菲奥（A. M. Fior）向塔塔里亚提出挑战，要求举行一次解三次方程的公开比赛，菲奥是著名数学家费罗（Scipiouedal Ferro）的得意门生，费罗大约在 1515 年左右解出了形如 $x^3+mx=n$ 类型的三次方程，并把方法秘密传给了菲奥。比赛于当年 5 月 22 日在米兰大教堂举行，双方各给出对方 30 道题。为了迎接这场挑战，塔塔里亚作了充分准备，他冥思苦想，终于在比赛前十天掌握了三次方程的解法，因而获得胜利。这种公开的解题竞赛无疑会引起人们对科学的注意和对科学研究的更大兴趣。

## 二、我国现代中学竞赛考试的发展

学校教育盛行之后，同伴间零散的竞赛型学习和竞赛型考试一直相伴相随。有一些竞赛活动可能是班级内部组织实施的，有一些是学校之间或者是地域型的。我国固定时间、全国范围的竞赛考试源自于由华罗庚先生倡导的全国数学竞赛。全国数学竞赛的第一届于 1956 年在北京、上海、天津、武汉四市举办，1964 年后因文化大革命停了 13 年，1979 年发展到全国 29 个省市联赛。上世纪 80 年代中期，我国各个学科陆续引入了奥林匹克科学竞赛，使之在全国的大范围内普遍推广。目前，全国每年有数百万的初、

高中学生参加各类学科竞赛，竞赛考试成为中学教学活动的一部分，是中学学生学业评价的一项重要补充，奥林匹克竞赛考试也成为仅次于中、高考的全国大范围考试活动。在理科奥林匹克竞赛考试的带动下，语文学科的作文竞赛，英语学科的英语知识竞赛也开始在较大范围内展开（如华东六省一市作文大赛、全国范围的新概念英语竞赛等）。

### 奥林匹克科学竞赛

奥林匹克科学系列竞赛包含各级奥林匹克数学、物理、化学、信息学（计算机）和生物学竞赛。

以化学奥林匹克竞赛为例，国际化学奥林匹克竞赛是化学资优中学生参加的最重要、最有影响的国际竞赛。其英文名称为 International Chemistry Olympiad，简写为 IChO，创始国是波兰、原捷克斯洛伐克和匈牙利。第一届 IChO 于 1968 年 6 月在布拉格举行，其后除 1971 年停办外每年举行一届，至 2008 年为第 40 届。虽然 IChO 是一种民间文化、学术活动，但在这 40 年中，它像滚雪球一样逐年扩大，覆盖地区愈来愈广。中国化学会自 1984 年以来，连续每年组织了全国高中学生化学竞赛活动。1987 年在全国高中学生化学竞赛的基础上，选拔成绩优异的高中学生参加了国际化学奥林匹克竞赛。截至 2007 年，我国选手已经参加 19 次赛事（第 37 届台湾赛事我国因故未能参加），共派出了 76 位选手，其中 52 人获得金牌，21 人获得银牌，3 人获得铜牌，他们为祖国和人民赢得了荣誉。

**表 13－1　我国首次参加国际各个学科奥林匹克竞赛时间**

| 生物 | 数学 | 物理 | 化学 | 信息学 |
|---|---|---|---|---|
| 1993 | 1985 | 1986 | 1987 | 1989 |

## 三、竞赛考试的积极作用

现代社会是一个竞争性社会，在学校中实施各类竞赛考试有利于学生形成竞争意识、激发学习兴趣、体会奋斗的艰辛历程并体验成功的喜悦。自中学开始的各类奥林匹克学科竞赛更是在学优生（也称为“资优生”）的能力培养上获得了积极的成果。竞赛的培训实施过程为参与竞赛的中学生的能力、品质的全面提高提供了一个很好的平台。

### （一）拓宽学生的知识领域，培养学生的学习兴趣

掌握基础知识是培养能力的前提，学科竞赛试题涉及到的内容广、层次深，甚至还涉及到许多新科学、新技术领域里的一些问题和方法，这些知识又是基础知识和方法的拓展和延伸。学生在准备竞赛过程中，不再囿于当前课程标准中的学科内容，在培训和辅导时教师会给学生提供较多的参考资料，这些资料开阔了学生的视野，使学生能够接触到更系统的学科前沿知识，提高了学习科学的兴趣，基础知识学好了，学生们才会产生灵感和悟性。通过系统的练习和思考，使学生对知识体系的理解和把握更为深刻，为创新思维铺垫了深厚的知识基础。由于竞赛活动的竞争性和层次性，参赛选手不断调整自己的学习目标，认识自我，体验成功，在向着新目标迈进的过程中获得了积极的人生体验。

### （二）锻炼学生的思维品质，铸造学生的创新精神

由于学科竞赛是学生自主参加，竞赛选手的学习过程基本上是靠选手自主学习完成的，选手通过广泛阅读资料，自主练习。学生为了达到目标，必须在阅读中读得细、读得深、读得准，自学时能自觉领会学习目标、准确把握重点。通过对竞赛选手的跟踪调查表明，竞赛选手的自学能力都得到了较大的提高，为其进一步深造提供了良好能力基础。

由于竞赛试题往往是在新情境下提供信息，考查选手的知识

和能力，对选手思维能力的考查要求较高，选手需要有较好的逆向思维、发散思维、聚敛思维、类比归纳思维等能力素养。在赛事训练中，选手还需要逐渐把握一些学科的专业思维，比如守恒思维、平衡思维等，有一些学科还要求选手掌握实验操作能力。在竞赛考试中，选手还需要形成灵敏、精确、整体、适应、实事求是等科学素养。

大胆假设、自我评价是成功的竞赛选手的基本素养，假设一方面是经验事实的总结，另一方面又是未知领域的新的探索。竞赛培训中，选手对学科内未知问题的探讨、对实验对象本质的探究，都是在学过的基础知识基础之上进行自我判断、自我评价的过程，这个过程也就是选手的创新思维、创新能力的形成过程。“不唯上”、“不唯知”是创新精神的前提，竞赛考试实施的过程能够很好的实现这一点。

## 第二节　中学竞赛考试的类型与方法

中学竞赛考试的形式多种多样，组织和实施方法具有比较固定的模式。教师在组织学生参加竞赛考试时，应依据一定的原则，合理挑选参赛考生，并正确看待竞赛考试，对学生进行合理的考试辅导，使竞赛考试原有的目标和价值得以体现。

### 一、中学竞赛考试的类型

竞赛考试活动按其最基本的组织形式来分，主要有活动竞赛和纸笔竞赛两种。

（一）*活动竞赛*

活动竞赛是以组织主题活动或者娱乐活动的方式举办的竞赛，又分为全员参与和选手参与型。全员参与型竞赛是所有成员都参加的竞赛活动，比如“查字典”比赛，活动在学生学完查字典并

掌握查字典的方法后进行，全班所有学生都参加。选手参与型竞赛是在每个团队中选出部分成员代表团队参加的比赛，得到的成绩属于所在团队。比如学校体育运动会，各个班级选出一定数量的选手参加，最终根据竞赛结果确定出优胜班级。值得注意的是，选手参与型竞赛是由代表选手直接参与比赛，而其他团队成员也都在某种层度上参与活动，所有成员都对竞赛成绩做出贡献。

（二）纸笔竞赛

一些主题竞赛活动模仿考试的方法采用纸笔测试方式进行。纸笔竞赛考试由于具有对选手考察全面准确、举办成本低等特点，在各级各类教学中采用较多。纸笔竞赛试题由竞赛组织单位在规定的范围（竞赛大纲）内命制，竞赛选手在大纲的指导下进行参赛准备，并在一定的地点（考场）规定时间内完成竞赛试题的考查。我国的奥林匹克学科竞赛主要是以纸笔测试的方式分级进行的，物理、生物、化学学科奥林匹克竞赛在全国范围内由初赛开始层级比赛，兼顾考查实验设计和操作能力。

也有一些竞赛综合两种模式，采用在一定主题下以活动竞赛和纸笔竞赛相结合的方式进行，比如英语知识应用能力大赛，一方面要书面考试，另一方面也要进行主题演讲。

## 二、组织竞赛考试的程序

以奥林匹克竞赛为例，我国的奥林匹克竞赛是各类竞赛考试中组织最为严密，体系最为健全的竞赛考试之一。各个学科的竞赛都在中国科协各相应学科竞赛委员会（以下简称竞委会）的领导下制定相应的章程，章程内容包含竞赛的目的、竞赛考试的性质，竞赛选手的产生培训选拔、竞赛的组织机构、竞赛试题的命制、评奖、国家队的产生等内容。以物理学科为例，中国物理学会常务理事会于1991年2月22日制定了《全国中学生物理竞赛章程》（以下简称《章程》），2004年11月30日通过该章程的修订

稿。《章程》是指导中学物理竞赛工作的纲领性文件，也是高中物理竞赛的考试大纲。

物理竞赛的组织和程序

（一）全国中学生物理竞赛每年举行一次，包括预赛、复赛和决赛。在校高中学生可向学校报名，经学校同意，由学校到地方竞委会指定的地点报名。凡报名参加全国中学生物理竞赛的学生均在地方竞委会指定的地点参加预赛。预赛由全国中学生物理竞赛命题组统一命题和制定评分标准，办公室统一制卷。各地方竞委会组织赛事和评定成绩。预赛满分为200分，竞赛时间为3小时。地方竞委会不得组织其它考试来确定学生参加预赛的资格。

（二）复赛包括理论和实验两部分。理论题由全国中学生物理竞赛命题组统一命题和制定评分标准，办公室统一制卷。理论考试满分为160分，时间为3小时。各地方竞委会组织赛事和评定成绩。复赛实验由地方竞委会命题和评定成绩，满分为40分，实验时间为3小时。复赛实验的日期、地点和组织办法由各地方竞委会根据实际情况自行决定。

参加复赛的学生由地方竞委会根据预赛成绩确定。参加复赛理论考试的人数不得少于本赛区一等奖名额的5倍。参加复赛实验考试人数不得少于本赛区一等奖名额的1.2倍。

（三）各地方竞委会根据学生复赛的总成绩（理论考试成绩和实验考试成绩之和）择优推荐3名学生参加决赛。对于在上届决赛中成绩较好的省（自治区、直辖市）给予奖励名额，凡有学生获一等奖者，一律奖励1个名额。在当年举行的国际物理奥林匹克竞赛中获金、银、铜奖的学生所在省（自治区、直辖市）每有1名学生获奖，就奖励1个名额。承办决赛的省（自治区、直辖市）参加决赛的名额可增加3名。

若参加决赛的最后一个名额有两名以上的学生总成绩相同，

则地方竞委会应根据他们的理论成绩高低择优确定一名；若理论成绩最高的学生有两名以上也相同，则地方竞委会可对理论成绩并列最高的学生以笔试的形式进行加试，选取成绩最好的1名。

决赛由全国中学生物理竞赛命题组命题和制定评分标准，决赛包括理论和实验两部分。竞赛时间各3小时。理论满分为140分，实验满分为60分。由组委会聘请高校教师阅卷评分。

由常委会聘请专家组成评奖组，由评奖组核审学生决赛成绩，提出获奖名单，最后由全国竞委会审议通过。

——摘自《全国中学生物理竞赛章程》

其他类型的竞赛考试程序一般也都是参照国际奥林匹克竞赛考试的模式，分为确定主题和对象、建立考试机构、设置考试范围、宣传培训、组织考试、考试评价6个环节。

## 三、组织学科竞赛考试的原则

### （一）选择合适的竞赛选手

竞赛考试，不是每个学生都适宜参加的活动，而是“学有余力，学有兴趣”的学生们参加活动。“学有余力”是强调首先要学好课内知识，在此基础上来学习课外知识；“学有兴趣”是指对学科有兴趣，正像有那么多“学有余力”的学生在学习画画、弹琴和唱歌一样。只有这两方面结合得好，才不会给学生增加学业负担。因此，选择什么样的选手参加活动，一方面要看选手的兴趣，另一方面要根据选手的能力确定竞赛的目标，总之要适合学生的身心健康成长。

### （二）平常心看待考试结果

竞赛必然是少数人获得胜利，但是并不意味着只有少数人有收获。在求得结果的过程中获得积极的体验是竞赛考试的初衷。学科竞赛考试是课程标准之外的课外活动，不能把它看得过重，

要尽可能的用“平常心”来对待它。更不应该赋予它那么多的功能：升学、选优、能不能拿诺贝尔奖等。一个学生取得好成绩，只能说明眼前，只能说明现在。如果一定要说对未来有什么影响的话，可能是在竞赛中获得的“成就感”会使学生对之后的学习、工作、生活增添一份自信心。

（三）难度要适宜

在开展竞赛活动中要秉持考试内容不超纲、不超前和“大众化、普及型”的组织原则。前者强调的是课内外的结合与一致，课内是基础、课外是补充；后者强调的是考试内容不宜过难，既不能让参与活动的学生感到高不可攀，还要让每个参与活动的不同层次的学生都有收获和提高，要坚持“普及与提高相结合，在普及的基础上提高”。

（四）适度辅导

教师在竞赛考试的组织和培训过程中，应着力于夯实知识基础和增强学生的自学能力，适度引导学生通过自己的努力完成竞赛的知识储备和能力养成，而不是包办代替或者实行题海战术。竞赛考试作为只有少数选手才能获得优胜的活动，在试题的设置上必然具有难度大、灵活性强的特征，竞赛考试不会也不应该成为通过机械训练才能有效应对的考试，选手只有在掌握基础知识并形成独立分析问题、解决问题的能力时才能胜出。

## 第三节　中学竞赛考试的发展趋势

纵观我国近年来中学竞赛考试的发展，可以说是“芝麻开花节节高”，各级各类中学竞赛考试在奥林匹克学科竞赛的积极影响和推动下，正逐步的发展和繁荣起来，为促进学校教育的发展发挥了重要作用。中学竞赛考试正呈现出如下特点和趋势：

## 一、竞赛形式更加丰富

中学的理科奥林匹克竞赛由于其操作规范有序，效果卓著，已经获得了全国广大教育界的普遍认可，也得到了社会各界的支持，在中学教育中得到了最大范围的推广。甚至与中学奥林匹克竞赛相衔接的小学奥赛也被社会热捧。理科奥林匹克竞赛的经验被推广到了其他学科及理科其他类型的竞赛中，如语文学科在全国范围内开展作文大赛；英语学科有英语能力大赛和新概念英语大赛；信息课程开展了机器人大赛、编程比赛等；数学除了奥林匹克竞赛外，陆续形成了希望杯（小学）数学竞赛、华罗庚杯（初中）数学竞赛、数学建模大赛等全国性的赛事。伴随着社会时事的变化，还有一些时事性的全国性赛事，如文史知识竞赛，“爱国主义”读书征文大赛等……

在平时的教学中，作为考试评价的补充，教师也越来越多地采用竞赛考试的方式增强学生学习的竞争性，甚至有的教师把竞赛用作课堂教学的一种策略来改变沉闷的课堂氛围，有效地激发了学生学习的兴趣。

## 二、竞赛目标更注重创新

伴随着历次的课程改革，中学竞赛考试的目标和培训模式也在变化改革中。

以中学化学竞赛为例：化学竞赛委员会每五年修订一次竞赛章程，在2006年的竞赛章程中提出的竞赛目的为“（1）普及化学基础知识，激励中学生接触化学发展的前沿，了解化学对科学技术、国民经济和人民生活以及社会发展的意义，学习化学家的思想方法和工作方法，以培养他们学习化学的兴趣爱好、创新意识、创新思维和初步的创新能力；（2）探索早期发现和培养优秀学生的思路、途径和方法；（3）促进化学教育教学新思想与新方法的

交流，推动大学与中学的化学教育教学改革，提高我国化学教育教学水平；（4）选拔参加一年一度的国际化学竞赛选手。”

而这之前的2000年竞赛章程中提出的竞赛目的为“（1）普及化学知识，鼓励中学生接触化学发展的前沿，了解化学对科学技术、社会经济和人民生活的意义、学习化学家的思想方法和工作方法，以激发他们学习化学的兴趣爱好和创造精神；（2）探索早期发现和培养优秀人才的思想、方法和途径；（3）促进化学教学新思想与新方法的交流，推动大学与中学的化学教学改革，提高我国化学教学水平；（4）选拔参加一年一度的国际化学竞赛的选手。”

对比这两次化学竞赛的竞赛章程，不难发现化学竞赛考试正是主动适应了我国新课程改革的需要，提出了更多的对创新意识、创新思维和创新能力的要求。

## 三、竞赛内容更注重能力考察

从近几年的全国学科竞赛试题来看，无论是试题的内容，还是试题的结构和题型都发生了深刻的变化，在很大程度上，学科竞赛已逐渐演化成智力竞赛（或能力竞赛）。因此，参赛选手单有概念、理论等知识的积累（哪怕这些知识面再广、再深）是无济于事的，重要的在于有灵敏的思维能力，有寻找、捕捉信息核对信息加工的能力，有很好的想象力和创造力，再加上良好的心理素质的配合，才能取胜。

### （一）试题内容范围更广

竞赛试题的内容范围进一步拓展，紧密联系生产和生活实际，关注社会热点问题，同时紧跟学科发展前沿。

### （二）竞赛试题结构更合理，题型更新

我国中学考试竞赛最开始的主流试题是构成题。这种题型是由题干和若干个问题组成，题干提供解题或形成试题的信息，问

题的提出和排列则是由命题人根据竞赛选手的知识和能力水平精心构筑而成的，故称为构成题。一般而言，问题按先易后难的顺序编排，最难的问题常常仅占该题总分的1/5左右，但该试题常常流于知识的罗列；试题设置或并列或递进，造成很大的随意性；而且大多数试题以考查知识的深浅度为主，不利于选拔创新人才。为解决上述问题，近年来考试竞赛试题中逐渐减少了构成题，取而代之的是一种新的主流试题——“科学猜谜题”，且权重越来越大。所谓“科学猜谜题”，它有别于通常意义上的猜谜游戏，其“谜面”是在试题中建构未知知识信息，猜谜人——考试竞赛选手的智力强弱表现在能否用已有的知识来理解这些信息，并对这些信息进行加工、分析、综合，最后创造性地形成谜底，即得出答案。一般“科学猜谜题”考察的是竞赛选手不知道的知识，而且是竞赛选手根据信息得出的“新知识”（有可能其知识细节对竞赛选手而言还不甚明了，但这些都不妨碍解题）。由于“科学猜谜题”的谜底经常出乎意料，它考查竞赛选手“推理破案”的能力，以及选手的“形成性”和“创造性”知识，往往用已有的模式来套反而得不出答案，因此很能考查竞赛选手的创造性思维的水平，既有利于考查选手思维的严密性、精确性、深刻性和全面性，同时也能做到试题的公正性，有利于选拔人才。该题型思考容量虽大，但应答书写少，也有利于评卷时减少误差。

这些变化使得偏重于考查学科知识的立意转变成为以考查竞赛选手能力为主的立意，即主要考查竞赛选手的创造性思维能力。试题尽可能使竞赛选手身处陌生情景，利用原有的知识基础，提取、加工、理解新情境下的信息，提出解决问题的方案、战略和策略，形成知识建构，发展应用能力，以达到考查竞赛选手学、识、才三者统一的水平。

## 四、竞赛选拔功能更强

我国中学竞赛考试一直以普及科学知识、激发青少年学习科

学兴趣、促进中学教学改革、探索发现科学人才的途径为目的，同时也有选拔大学免试保送生及参加国际奥林匹克竞赛选手的功能。从选拔功能上看，由于加大了对选手能力的考察力度，考试的题型、结构等都做出了一定变化，赛题的选拔功能更强了，有助于吸引和选拔更多资优学生参与竞赛。

中学竞赛考试呈现的上述特点和趋势对参与竞赛的教师和学生们提出了更高的要求，无论是从知识基础、能力品质，还是身心素质上，都必须进行科学、系统的训练，才能适应不断向精、尖发展的竞赛考试的考验。同时，竞赛考试体现出的更加注重对创新意识和综合能力的考察等趋势，将对中学考试的发展和人才的培养起到积极的影响。

# 第十四章　国外教育考试评价改革评述

教育考试历来是世界各国普遍关注的重大课题，由于各国的社会条件不同、教育制度和传统不同，各国的教育考试制度呈现出不同特点。近年来，随着社会发展及人才需求的不断变化，世界各国的教育考试也在不断的改革发展中，尤其是大学入学招生考试的改革，呈现出百花齐放的局面。虽然各国做法不尽相同，但整体趋势是相近的。

## 第一节　国外教育考试评价改革的发展与特征

国外教育考试有着悠久的发展历程，至今已经历了几次重大的变革，出现过几代不同性质的教育考试评价模式。

### 一、国外教育考试制度的发展历程

（一）国外教育考试评价的演变

关于国外教育考试评价演变，国内外学者的描述不尽相同，这里介绍教育考试评价的“四代说”[①]。

1. 以笔试为标志的前教育评价（18 世纪初 ~ 19 世纪中叶）

十八世纪以前，西方学校教育测量主要是采用口试的方法。

① 史耀芳．国外学校教育评价的历史沿革［J］．外国中小学教育．1997，（4）：20－23.

1702 年，英国剑桥大学为了改革评价方法，最早采用笔试取代口试作为评价的主要方法，开西方学校笔试之先河。笔试的引入使考试方法的客观性和可靠性都有较大的提高。但是，当时的笔试无论从命题上看还是从评分的过程分析，主观随意性都很大，不能很有效地反映和评定学生的学力水平，而且也不便于比较。

2. 以测量为标志的第一代教育评价（19 世纪后期～20 世纪 20 年代）

英国的心理学家高尔顿通过对个体差异的长期研究，于 1869 年发表了《遗传的天才》一书，揭开了教育测量的序幕。1904 年，美国的心理学家、教育家桑代克发表了《心理与社会测量导论》一书，系统地介绍了统计方法和编制测验的基本原理，奠定了测量在教育评价中的绝对地位，使测量作为教育评价的方式兴盛起来。以后的几年中，诸如比纳——西蒙智力量表、斯坦福——比纳量表等心理测验的日趋定型，对教育测量的标准化也产生了极其深刻的影响。这一时期教育测验的发展表现出以下几个特点：（1）单科测验向综合测验发展；（2）由小学向其他学校发展；（3）随着心理测验的发展，教育测量的客观化、标准化受到了极大重视；（4）心理与教育测量的理论得到了充分发展，并日趋成熟和完善。为此，这一时期被称为“测量”时期。

3. 以描述为标志的第二代教育评价（20 世纪 30～40 年代）

1933～1940 年，在进步主义教育协会的领导下，对美国七所大学、三十多所中学进行了为期八年的教育改革实验研究（即：“八年研究”），研究结果表明当时的课程与测验是教科书中心主义，测验内容只是记诵教材的知识内容，并不能真正反映学生全面的发展。为此，他们提出了一套以全面发展人的才能为核心和依据的课程编制及测验编制的原则，对教育后学生成绩的考试方法进行了积极探索，明确指出：（1）评价应该是一个过程，而不仅仅是几个测验；（2）评价过程不仅要报告学生的成绩，更要描

述教育目标与教育结果的一致程序，从而发现问题，改进课程教材和教育教学计划及方法。

“八年研究”实际上宣告测验运动的终结，使人们在思想上和行动上接受了采用教育评价方法来描述教育效果这一观点，形成了一个以“描述”为特征的评价时代。

4. 以判断为标志的第三代教育评价（20 世纪 50 年代末、60 年代初～70 年代）

1957 年，前苏联第一颗人造卫星上天使美国朝野大为震惊，促使美国投入大量的人力和财力进行大规模的教育改革，从而对教育评价也提出了新的要求。1967 年，美国教育学家斯泰克发表了《评价的面貌》一文，肯定了判断是评价的两大基本活动之一，同时提出了一个完整的、包含描述与判断二个方面的评价模式。这样，判断就成了第三代教育评价的标志。正如美国著名教育评价专家顾巴所说的：“60 年代或者更确切地说 1967 年以后，判断成了第三代评价的标记。”

5. 以回应和协商为标志的第四代教育评价（20 世纪 80 年代至今）

第四代教育评价的创立者、美国著名教育评价专家顾巴和林肯认为第四代评价为了打破以往评价中“管理主义的倾向”，使评价成为各有关方面共同的事情，评价的起点就应该是对有关方面评价要求的“回应”，并明确指出：第四代评价的模式就建立在建构性探究方法的基础上，同时强调：探究的过程本身也是一个不断协商的过程，一个不断发现新问题、不断验证新的共识的过程。

从前教育评价到第四代评价的历史演变中，我们可以发现并总结出国外学校教育考试评价的发展有如下特征：

（1）从评价的目的作用上看，已由单纯地服务于人才选拔，发展到服务于人才培养、教育改革，进而提高教育的质量和效益；

（2）从评价的步骤程序上看，更加注重自我评价的作用，并

划归为预评价阶段，便于更全面地收集信息，也有利于改进教育工作；

(3) 从评价的内容范围上看，由单纯评价学生的学业成绩，发展到评价学生与评价学校、教师、基层教育行政管理部门相结合；

(4) 从评价的计量方法上看，由简单的定性或定量方法发展到注意定性与定量相结合、客观与主观相结合，由简单的教育质量评价发展到教育质量与教育效益评价相结合。

(二) 国外教育考试评价模式的改革

在教育考试的发展历史上，曾出现过许多具有不同背景、不同价值取向、不同目的与方法的教育考试模式。本书在这里主要介绍以教育考试的价值核心为标准划分的四种评价模式①。

1. 目标达成模式

教育考试的目标模式将考试看作是学生学习的结果与预定教学目标相对照的过程。其隐喻是：学校即工厂。学生是等待加工的产品，教师是负责产品质量检验的技师，教学系统就是一条生产流水线。目标模式的代表人物是美国的教育家拉尔夫·泰勒。

教育考试的目标模式强调教学的效率化、标准化、模式化。目标取向的评价模式受“科技理性”的支配，追求评价的科学性与客观化，以自然科学范式为理论基础，基本方法是量化方法，具体方式主要是终结性的评价。历史地看，目标模式加速了考试科学化进程，它简单易行，容易操作，并已形成了一套完备的评价程序与评价方法。时至今日，目标模式仍然在学业评价的实践中发挥着重要作用。

2. 诊断反馈模式

考试的诊断模式将学业评价看作是诊断与改进教和学的过程。

① 刘志耀．基于教育隐喻的学生学业成就评价模式探析［J］．太原大学教育学院学报，2007，(25) 3：30－33.

其隐喻是：教室如诊所。教师的工作就是为学生的学习查找“病因”、然后“对症下药”。诊断模式的代表人物是美国的心理学家布卢姆。

诊断模式下学业评价的任务就是通过测验，查找教师的教学与学生学习中的困难与问题，然后将这些信息反馈给教师与学生，并在此基础上改善教师的教与学生的学。诊断模式受“工具理性”支配，它强调通过科学化、客观化的评价（主要是测验）来改善教与学，使用的是量化方法和科学范式。该模式实际使用的是以查找问题为手段、以纠错矫正为目的形成性评价的方法。

3. 过程情境模式

考试的过程模式，是将学生学习的全部过程都纳人评价范围的学业评价范式。其隐喻是：教学如旅行。课程被看作是学生旅行的线路，教师既是有丰富经验的导游，也是学生旅游的伴侣。过程模式的代表人物是美国的斯克瑞文和英国的斯腾豪斯。

过程模式下评价者特别珍视每一个学生的个体教育经验，以及教育效果的不可预测性。这样一来，评价者所关注的，就不仅是有关教与学的具体数据与最终结果，而且关注教学过程对于每一个学生、以及不同学生群体的不同影响和效果。总体说来，学业评价的过程模式同关注学习结果一样地关注学习过程，它受到“实践理性”的支配，强调学习过程本身的价值，并开始意识到评价是一种价值判断的过程。与前两种评价模式相比，过程模式对于学生在学习过程中的行为表现给予更多的关注，也对学生在学习过程中的主体性、创造性给予更多的尊重。

4. 主体参照模式

考试的主体模式则将学业评价视为评价者与被评价者，具体来说，就是教师、学生、家长等共同构建意义的过程。其隐喻是：校园如花园。教师是园丁，学生就是花朵。园丁的工作就是给它们浇水、施肥、修枝整叶，为它们的正常生长提供合适的阳光、

水分与养料。

在主体模式下教师该做的事，不再是教育目标的预设、知识内容的灌输以及行为方式的规范，而更多的是提供信息咨询与服务。在学业评价的本质问题上，主体模式认为，评价是一种价值判断的过程，但强调价值的多元性，评价者与被评价者、教师与学生在评价过程中是一种交互主体的关系，评价过程是一种民主参与、共同协商和平等交往的过程。所以，主体模式强调使用自我参照、自我接受式的评价方式。

应当指出的是，任何一种学业评价的主体模式都仅仅代表着学业评价的一种可能的发展方向，但它不是也不可能是当今学业评价的主流。即使在将来，任何一种模式也不可能成为学业评价模式体系的主体，更不可能成为全部。因为它主要关注人并服务于人的发展，而社会的发展需要则必须发挥其他评价模式的功能。所以从这个意义上说，在任何时候，我们对于评价模式的选择都将是复合的而不是单一的。学业评价的主体模式应该以一种理念的方式春雨润物般地渗透到学业评价的活动中去，而不应该以制度的名义和强制性的方式付诸学业评价的实践。

## 二、部分国家教育考试改革简介

国外一些发达国家对学生的考核方式，除考试外，更注重对学生日常的考核，作为评价用具的考试更加多样化。从单纯的笔试转向综合考查，有口头报告、作文、调查、实际操作等形式，在考核中往往是多样形式的综合运用。下面介绍国外部分国家的教育考试改革的特征及趋势。

### （一）美国[①]

1983 年，美国公布了对教育进行尖锐批评的文件——《国家

① 曾雄军．当代美国基础教育改革中的三种策略［J］．外国教育研究，2002，29（11）：33－36.

在危机中》，这份报告列举了美国教育的种种败绩，使美国公众对教育的现状深为不满和担忧，普遍要求改革学校教育。为此 1986 年，美国决定强化国家、州一级对中小学校的管理，实施统一的国家或州标准化考试。此项措施主张运用全国或者州范围内统一的“标准化考试”对教学结果进行评价，评定学校等级以及选拔优秀学生——决定哪些学校的教学质量需要改进，哪些教师符合学校要求以及哪些学生能够从中学毕业等。

目前，由于美国以整个高中阶段总成绩为主要升学依据的考试制度为素质教育提供了充分的时间和空间，使得中小学的素质教育真正落到了实处，所以美国中小学生虽然学习任务较重，但他们并不感到负担沉重。美国学校几乎每门课程结束时都要进行一次期末考试，大部分课程安排期中考试，课程进行中间还有小测验，所有考试的目的完全在于确定学生是否真正完成规定的学习任务并掌握了授课内容，考查学生对所学课程的理解程度，考核的侧重点是学生分析问题和解决问题的能力。学生不必死记硬背教材的条条框框，只要尽可能发挥各自的想象力和创造性思维，问题的答案只要能言之有理或自圆其说都可以得到理想的成绩。考试根据不同专业、不同课程的特点采用闭卷、开卷、口试、答辩及实际操作能力测试等多种形式。在学生学业成绩的评价方面，美国学校建立了科学的综合的学生学业评价体系，采用等级评分、固定成绩评分和成绩分布曲线评分等相结合的方法，弱化考试的总结评价功能，这样可以更好的鼓励学生学好知识。

（二）英国

1988 年，英国公布了《1988 年教育改革法》，它是《1944 年教育法》公布以来最重要的教育法案，标志着英国教育制度的巨大变革。它扩大了中央政府的权力；赋予家长更大的选择权力，并要求家长承担更多的职责，要求地方教育当局下属的所有中、小学推行中央政府指导的国家课程和学生评定计划；削弱地方教

育当局的作用，让学生享受更大的自治权，并建立享有更大权力的理事会；创立一种新型的、由教育科学部直接控制的公立学校[①]。

2004年，为了解与检查英国学校教师对学生进行评价的具体情况，英国政府投入了45万英镑的资金，这一举动，带来了英国国家统一课程考试形式的改变。目前，英国13个地方当局的74所学校的教师正在试行一个《重要的3阶段计划》，试行这一计划的每所学校中有两名教师正在使用一个被称为“监督学生进步”的新的评价制度[②]。“监督学生进步”这一新的评价制度蕴涵了评价的内容与方法。被挑选出来的学生将在每个学期参加一次指定材料的阅读与写作活动，并在规定时间内默写出由课程和资格认证机构要求的知识点，然后，教师将依据学生所默写的知识点来给学生进行打分。而打分的原则与规定由相关的考试监督机构制定。这些规定很复杂，目的是确认学生的学习优势和劣势，帮助家长进一步了解子女的学习情况，并为教师确定未来的教学目标提供便利条件。

英国的教育考试发展总体来说有三个特征和趋势[③]：

1. 向统一考试的方向发展。目的是为学生提供平等的机会，统一考试，统一标准，也便于管理和评价学校的工作，同时利于学校间的竞争。

2. 重视学生平时成绩，在证书考试中加入一定比例的学科作业成绩。这种做法可以减轻学生考试的压力，同时增强学生平时学习的动力。在义务教育阶段推行平时学业成绩的考核和记载制

① 王斌华．国家课程和学业评定——英国教育改革重大举措之一［J］．外国教育展望，1994，（3）：8－15.

② 杨聪聪．美、英、日中小学学业评价前沿探究［J］．世界教育信息．2007，（1）：27.

③ 杨春增．国外中小学学业成绩评定办法［J］．天津市教科院学报，2001，（6）：48－50.

度，以利于经常地、全面地评价学生在各个学段、各个方面的成绩。

3. 从重视终结性评价转向重视形成性评价。1988 年英国教育改革之后，英国在义务教育阶段划分为4 个基本学段，结束时都要进行考试。把过去的终结性评价转化为形成性、诊断性、终结性和评价性相结合的评价，以此促进学校教学质量的提高。

（三）澳大利亚[1]

澳大利亚现行的考试制度是传统与现代考试制度的结合，反映了当代社会经济发展对教育的需求与挑战，在满足学生以及社会各方面对教育的需求方面，不断进行调整与变革，逐步建立起适应各州或各地区实际的教育考试制度。

澳大利亚小学很少对学生进行考试，而到了中学就有了严格的考试制度。除了平时的阶段考试外，在 10 年级结束时学生要参加全州统一举行的毕业会考。考试成绩基于学校平时成绩以及由全州统一标准的数学、英语和自然科学会考成绩综合评估而定。12 年级结束时还需要参加州里的高中毕业会考，以获得高中毕业证书。

但是，近年来各州已经不再把考试作为考核学生学习成绩的惟一标准。任课教师对学生的学业评价，不仅仅关注相关测验或考试的分数，更注重学生学习过程的学业评价。在这个学业评价体系的评价项目中，还包括许多对学生在学习过程、问题解决过程中的表现等方面的评价。而这些具体的评价一般是围绕着学生在每个学习阶段的学习成果展开和进行的。

不难看出，澳大利亚的中学考试制度以人为本，充分考虑到尊重学生的不同兴趣特点，最大限度地调动学生的学习积极性，为学生提供广阔的学科知识发展空间，有利于学生未来的发展。

---

1 牛道生．澳大利亚中学考试制度述评［J］．湖北招生考试，2003，（4）：55－57.

（四）法国

在法国，对学生的评价从重视考试成绩转向对学生的全面考查，评分方式从注重分数转向注重实质性的评价。法国中学对学生学习情况的考评与小学基本相同，采取“持续观察”的方法。教师与学生家长之间的日常联络是通过联络卡进行的，联络卡上除印有学校的内部规章外，还记载着学生平日上课的出席情况、学习情况、行为表现等内容，以便学校能将学生的在校情况及时通报给学生家长。每学期期中和期末，学校将学生手册寄给学生家长，手册上详细记载着学生各门课程的学习情况和成绩以及教师的评语。

（五）日本

日本的学业评价一般是教师根据文部省颁布的学习指导要领中的学习目标分类，对学生的学习成果进行评价。越来越多的教师为了获得更多的学生信息以调节教学，更多地关注发展性目标和情意性目标的评价，因而比较多地注重采用现场观察、行为记录等定性评价的方法。而除了完成教育部制定的《学习指导要领》中的各项内容的评价工作外，教师在日常教学中还要根据不同的需要采取多种形式进行评价，如论文体测验、实验操作的观察、实际生活的观察、测验对不同情境的反应、自我评价、问卷、面试、事例研究等。

综上所述，国外考试评价改革总体呈现下列一些趋势[①]：

1. 限制统考次数，减少考试难度，强化课程作业。

2. 对于考试手段的把握，基本上呈现出小学阶段较宽松、初中阶段略有扳紧、高中阶段相对绷紧的趋势。

3. 课业考评方法多元化，教学与评价整合化，尤其是表现性

① 黄光扬．关于基础教育考试评价改革若干问题的探讨［J］．课程．教材．教法．2004，（5）：87－92.

测验和实验技能教学考试受到高度重视与广泛采用。

4. 在对学生学习成绩和学生素质发展评价时，大量使用观察表现的等级评定量表，同时普遍重视学校记录和学生成绩书面报告单。

5. 评价学生强调个别适应性，参与评价的人员多元化和互动化。

6. 考试评价结合学生年龄和学习内容特点。

## 第二节 国外大学入学招生考试改革

大学入学考试是教育考试制度的重要组成部分，并常常是教育考试改革的突破口。近年来，美、英、法等国都进行了一系列的高考改革，他们在招生时不仅重视学生的分数，而且对学生各方面的能力都给予了充分考虑。

### 一、国外大学招生制度改革发展趋势

目前世界部分国家大学招生考试制度主要有三种模式：综合选拔制、证书制和高考制。综合选拔制是资格认定和基准淘汰相结合的制度，以美国为代表，招生政策灵活，考试形式、录取标准和招生方式多样化；证书制以英国为代表，高校一般不组织入学考试，而是根据入学申请者参加高水平普通教育证书考试的成绩来选拔；高考是竞争选拔的方式，亚洲国家采用者居多，除中国外，以日本为代表。尽管各国模式不一，但在招生制度的改革方面，存在许多共同点。

1. 大学入学招生考试由分散逐渐走向统一

传统的大学入学形式复杂多样，考试内容、考试时间等的不统一，造成大学招生的不公平、不公正现象时有发生。为了使大学入学考试更规范化、合理化、人性化，考试的结果更具有可比

性，大学入学招生考试正由分散逐步走向统一。

2. 考生与高校双方自主权加大

在考试科目和考试时间等的选择上，改革也给与考生很大的自主权，这样有利于培养考生的个性特征，对学生的长远发展有利。

除了考生，各招生学校也拥有很大的自由空间。部分国家的大学依据资格考试，自主决定考核科目、录取标准等，在自行组织的个别考试中享有完全的自主权，从而录取适合自己学校的学生，有利于大学自身的发展。

3. 统一考试与大学单独考核相结合

大学进行单独考核，能较好的将高校的理念与特色和学生的志向与特点结合起来，也易于考查学生的综合素质和能力。但单独考核对比性差，公平性难以确保，考试成本也较高。为了弥补这些不足，统一考试愈发受到各国的重视。目前的趋势是统一考试与各大学举行的单独考试相结合。

4. 考试管理由政府行为向社会化经营方向发展

总体上看，考试管理社会化已逐步形成几个模式：（1）大学附属的独立考试机构，如美国普林斯顿教育考试中心。（2）校外的非盈利性教育考试机构，如英国主办考试的就不是政府部门，而是一些与高等学校有紧密联系的非盈利机构，但主办机构必须经政府教育部门批准。（3）社会经营性的考试公司，即考试由民间机构以有偿服务的方式举行。

5. 考生考试机会增多，选择空间加大

考试机会的增多是高考改革为了适应终身学习社会的重要内容。包括同一考生和不同考生两个方面：

对同一考生来说，传统高考具有一考定终身的弊端，为此一些国家的大学入学考试在一年内进行多次考试，这样增大了考生成功的几率，也增大了学生的选择空间；对不同的考生来说，高

考既要面向应届的高中毕业生，同时也要面向其他各种社会成员，很多国家已采取各种措施，以给不同的考生更多的考试机会和选择空间。

6. 选拔方式多样化，既重视高考成绩，又重视中学阶段的学业成绩

大学招生的选拔方式更加多样化，不仅重视学生的统一入学考试成绩，还重视学生中学阶段的学业成绩，包括校长推荐信、教师评语、个性特长、社会服务等内容的综合评价。这样既有利于学生综合素质的培养和多种能力的发展，也有利于学校更加全面的选拔人才。

7. 考试内容注重知识的考核，更注重学生综合素质和能力的考察

在考试内容方面，不仅重视考生中学阶段各方面知识的考核，更注重考生综合知识和能力的考察。考试内容很注重考生创新思维、口头表达等各方面能力的考查，这有利的引导了中学培养学生的理念和方向的转变。

## 二、部分国家大学入学招生考试改革简介

### （一）美国：注重学习潜能，测试形式多样

美国大学入学要求最重要的4个因素是：大学预修课程（AP课程）的成绩、全美统一的入学考试（SAT I/II、ACT）、高中全部课程的成绩和班级排名。次重要的4个因素有：入学短文、选课指导教师的推荐，任课教师的推荐和口试。最后，学生在课外活动、校外工作中的表现、州毕业考试、学科考试等因素也占据了一定的分量，而且还呈现上升的趋势。美国的大学入学考试制度主要表现在以下几个方面：

1. 高中学术性科目的平时成绩。评价学生平时成绩，不仅仅是期末的考试，还取决于平时的各种项目表现，如实验、写作、

论文等，以及课堂上的参与程度。

2. 大学预修（AP）课程的统考成绩。AP 课程（Advanced Placement Program）是大学水平的课程，高中学生修读此类课程，既有助于申请大学入学，又可以免修大学的有关课程。AP 科目种类很多，约有三四十门（AP 课程考试的题型分为客观题和主观题两大类）。

3. SAT 和 ACT 统考。美国也有类似我国的统一高考，但是它只占大学入学要求的一小部分。美国的“高考”包括学术性向考试（SAT I）、学业成绩考试（SAT II）以及美国高等学校考试（ACT）3 种类型：

（1）学术性向考试（SAT I）是采用最为广泛的高校入学考试，它主要测验学生的综合学习能力，并显示学生在全国大学考生中的成绩排名先后。该考试只有英语和数学两门，每门800 分为满分。此种考试每年举行6 次，在全美国和世界各地设了数百个考试点，学生可以自由选择合适的地点和时间参加考试。

（2）学业成绩考试（SAT II）共有 15 门考试科目，包括写作、文学、美国历史与社会研究、数学、物理、化学等，学生可以任选其中的3 门。有的院校将此种考试成绩作为录取学生和对学生入学后因材施教及指导专业选择的重要参考。

（3）美国高等学校考试（ACT）包括英语、数学、自然科学和社会研究4 个部分，主要考查学生的基本知识和基本技能。此项考试与学术性向考试颇为相似，且在许多情况下可互相取代，每年也举行6 次，考生可自由选择时间。

不同的大学，对于3 种考试的要求是不一样的。如哈佛大学规定申请者均需要参加 SAT I 或者 ACT 考试，以及参加3 门科目之 SAT II 考试（其中的英语熟练考试是必考的）。

此外，高中推荐、开卷的入学短文、面试和课外活动和校外工作的平时表现都可作为大学入学考试成绩的参考。

（二）英国：重视考查学生技能和实践能力

根据英国政府规定，在义务教育结束时，要为16岁的初中毕业生提供一种“普通中学教育证书考试”（General Certificate of Secondary Education，简称GCSE）①。考试由英国教育部门统一命题，考试内容为英国《1988年教育改革法》中规定的义务教育阶段公立学校应开设的必修课，通常要考8到10门课程，其中英语、科学和综合理科三门课是每个学生都必须考的科目。考试时间全国统一，考前学生可复习教育部门指定的复习大纲，大纲分为A、B两类，但差别不大，考试内容主要是后两年学的课程，其目的是为了减轻学生负担。

GCSE考试还设有“课程作业”评定项目，重点考查学生的设计、操作、观察、口头表达能力和组织能力。《GCSE国家标准》规定所有的考试设计都应适当的区分度，即“能力区分度”评定，这主要体现在各种考试是由不同难度水平的试卷组成的，一般分为普通和扩充两套试卷，考生可根据自己的实际水平选做某种试卷。通过GCSE考试的学生如果想进大学深造，还必须在高中学习两年，这叫延续教育，这除了上大学外，还可为未来的就业作准备。读完高二后，学生必须参加“普通教育高级考试”（A－Level）考试的课程有英语语言、数学、生物、化学、地理等，学生选修的课程也要考试。考试的评分标准分8档，从A、B、C一直到G，零分则用U表示。

另外，英国的考试制度非常重视学生独立思考能力的发展，许多课程都要求学生写论文，如历史论文成绩占40%，其它知识占60%。论文一般为开卷考试。因此，英国GCSE和A－Level考试很注重引导学校培养学生实践及独立思考问题的能力，这样培

① 杨光富．当今美英法日四国高考制度［J］．外国中小学教育，2002，（2）：41－43.

养出来的学生会更好地适应以后的大学学习和生活。

（三）俄罗斯：两次考试转为全国统考

随着俄罗斯经济体制的改革，教育领域也全面实施改革。2004年起在俄罗斯全国范围内实施全国统一考试，取代了原有的中学毕业考试和大学入学考试。

考试的基本科目包括数学、俄语、文学、物理、化学、生物、地理、俄罗斯历史、社会常识、外语，根据所报专业进行相关科目的考试。

考试内容以俄罗斯教育部制定的普通教育科目示范大纲为基础，不能超出大纲规定的范围。所有科目的测试题（除外语）都由难易程度不同的各种题型构成。题型分三部分，第一部分为客观性试题；第二部分是介于主观性试题和客观性试题之间的试题，叫限制性试题（这类试题包括简答题、填充题）；第三部分是主观性试题，包括论述题、证明题、作文题等。外语试题由五部分构成，四部分笔试（听力、阅读、词汇与语法、作文），一部分口试。按国家统一规定的基础科目对应试者进行基础知识和基本技能的考查，口试范围缩减，笔试范围增加，这种考试对毕业生来说更加公平、公正，对保证新生质量具有重要意义。

俄罗斯原有的“五分制”评分标准不够细，随意性也较大。因此统一考试的结果设置了两个评价体系：五分制和百分制，即按百分制进行评定后再按联邦的规定转换成五分制。两种评分标准的统一是在同一分度上完成各种试题的分数进行折合的基础上达成的，这些试题能够真实反映出毕业生的知识水平。

此外，为了照顾农村和偏远地区的考生，俄罗斯把运用远程教育手段和因特网条件放在了重要位置，力求使全国统一考试变得更加公平、公正。

由此看来，俄罗斯的国家统一大学入学考试无论是在规模上、题型结构上，还是在评分标准上、公平程度上都要优于原有的考

试制度，并且朝着信息化、网络化方向迈进。

（四）澳大利亚：高考和平时成绩各参半

澳大利亚阶段考生入大学的成绩由两大部分组成，一部分是高考成绩（又称高中联考)，另一部分是平时成绩各占50%；平时成绩由三部分组成，模拟统考占20%，12年级二次测验（每学期一次）占15%，学科特色分（比如物理化学是实验）占15%。这在很大程度上避免了一次失误全盘皆输的局面①。

澳大利亚的高考考试有十八个大类，每类又下设若干科目。例如外语是一类，这类中有许多语种。18类总共有几十个科目供考生选择，而且每一科的考试与中学课程设置完全匹配，根据各科的学时数、内容多少、难易程度等因素，每一科都分出级别，级别越高，难度越大，分值也就越大。高校各专业则在招生简章中规定其必考科目和级别。大学按考生等级分择优录取，等级分表示该考生在所有考生中的排名位置，这就解决了分科分级带来的横向比较选拔的技术问题。在这种考试设计下，学生有相当大的选择余地，可以根据自己的所长再结合心仪的学校的要求，去选择所考的科目以及所考科目的级别。

澳大利亚的高考持续1个月之久，从每年的10月下旬一直考到11月下旬，这样不会因考试时间重合而让考生错过选择机会，在时间上充分保证了考生的选择。

（五）法国："双轨制"选拔严格

法国高考制度实行双轨制，简单来讲就是法国高中毕业举行高中会考，通过的学生可获得业士学位。大部分的业士学位获得者向各大学申请入学资格。而剩下的一小部分的业士则会选择进入大学预科学习一至二年，然后参加极其严格的竞争淘汰式考试，力争进入为法国培养精英人才的工程师学校②。

---

① 阿彬．发达国家高考制度比较［J］．理论与当代，2003，(9)：35－36.

② 刘学伟．法国高考制度改革简介［J］．教育与职业，2003，(11)：56－60.

法国的高中会考（简称 BAC），它属于全国性的考试，每年举行一次，由国家统一出题。由于法国的高中是实行分科教学的，高中各科和各种方向的考试科目和内容是不同的，因此高中毕业会考分为普通类和技术类两大会考。会考内容依据国家规定的高三课程纲要而定。因为考试科目的不同，每一科在评分时加权的多少由一个系数来表示。法国为 20 分制，总成绩如果平均分达到或超过 10 分者就算通过了；得到 8 或 9 分者有资格参加第二次的考试，即补考；8 分以下算未通过。总的说来法国的中学毕业会考比较严格，质量也较高，其及格率一直控制在三分之二左右。

另外，法国高校为了提前掌握生源情况，在招生中采用了入学预注册制度，即提前让学生登记志愿，每人至多登记 3 个志愿，但仅含 2 个科系及两所大学。

（六）日本：扩大高校招生自主权

近年来，日本文部省和各高校对高考方式进行了积极的改革，其主要措施有高校选拔学生的方式多样化、扩大推荐生的招收规模、实行“引入考试”（AO）的招生方式等。

日本高考分两个阶段，成绩优劣决定人生命运。日本的高考本来是各高校自己单独组织的，1976 年经文部省批准，成立大学入学考试中心，负责全国的高校的入学考试。但是现在日本的高考把全国统考和学校单独组织考试两者结合起来，因此，高考全过程分为两个阶段：第一个阶段是国立、公立和部分私立大学第一次全国统考，时间为每年的元月份。由大学入学考试中心统一组织。主要考中学必修的科目，主要包括基础学科（日语、英语、数学），理科（物理、地理、化学、生物），社会学科。通过这次统考的学生，还必须参加第二个阶段的考试，即各校根据自己的专业特点和要求，进行的专业考试，主要考其学术性问题，出题范围主要是中学的选修科目，时间一般在每年的 2 月中旬至 3 月底，学生可以根据自己第一次考试的成绩参加各校符合自己志愿

的考试，其它没有参加统一考试的私立大学均自行招生。

（七）德国：先笔试后口试

德国的大学招生体系具有多主体和较大的自主权的特点，参加考试的学生在高中的第四学期初，由校长根据学生的平时成绩来确定学生是否有权参加毕业资格考试。考生要在完全中学的12～13学级阶段至少各完成4个半年考试学科的学程，并获得合格的成绩，才可准予参加考试。此外，考试必须证明已学过的基础学程达到完全中学高级阶段规定的规模。

考试共分四类科目，采用笔试和口试两种方式。考生可以在四类科目中选择四门作为考试科目，其中前三门为笔试，第四门为口试。考试于第四年第二学期进行，从学期初的2、3月份开始，先进行笔试，再进行口试，最后用四周左右时间评分①。

（八）韩国："大学修学能力考试"亮点多

韩国政府为推进基础教育从过去的应试教育向素质教育、创新教育发展，引进了美国的SAT大学入学资格考试制度，将大学统考改为"大学修学能力考试"，注重学生的人格与能力培养②。具体措施有简化高考科目、压缩高考时间、减少考题量等。此外韩国政府还从法律上规定各大学不得再进行正规的二次考试，只能采用小论文、面试等小测验录取。高中综合成绩在大学录取中也发挥着重要作用，有特长的学生可以通过"特别选考"进入大学。

2002年，韩国政府对大学修学能力考试进行分析后，又提出了从2005年起进一步增加学生选择机会的更加合理、完善的考试方案，即数理科目分成数理A和数理B两种，社会和科学探究附

① 牛学敏，王后雄．德国大学招生与考试制度改革述评［J］．中国考试（研究版），2007，(7)：53－58.

② 李水山．韩国招生考试制度的改革沿革与发展方向［J］．基础教育参考，2005，(2)：14－15.

加职业探究，第二外语改为第二外语加汉文，社会和科学探究又分成若干增加所属某一课程比重的考试科目，供学生选择其中一门，增加了资格考试的多元化、选择性、科学性和合理性。

## 第三节　国外考试改革对我国的启示

由于历史、社会以及文化等方面的原因，总的来说外国中学的考试制度都体现了灵活性、开放性和多元性的特点，很大程度上呈现了崇尚自由、鼓励创新、勇于实践和彰显个性的精神实质。国外中学考试的经验和发展趋势对我国中学考试评价制度改革具有以下启发和借鉴的意义。

### 一、改革考试内容

考试内容应当作为测试、衡量和提升学生思维能力的有机组成部分，应注重考查学生运用知识的能力；考试内容要富于时代气息，应与社会现实问题衔接，考核学生多样性知识的应用，加大非智力因素考核力度，通过标准答案的多样性、灵活性为学生想象力的发展创设一种弹性环境。

### 二、改革考试方式

考试方式上要具有多样性和开放性，学校应根据自身条件以及学科性质考查学生综合素质，采用闭卷、开卷、小论文与答辩、社会调查报告、实验设计、试验创作等多元化考试形式，允许学生根据自我个性和实际在多样化考试中自主选择。同时，应建立综合考查机制，将其平时表现与平时成绩综合到终结性评价之中。考试标准要有决定性标准和个别化标准，对学生进行横向和纵向比较，明确考生素质优劣及潜力所在，为学生提供帮助。此外，为了提升人才的国际竞争力，很多发达国家的高校把口试作为笔

试的辅助性考核手段在一些课程中做了硬性规定。口试不仅可以有效防止考试作弊现象的发生，更加真实地反映出考生水平，而且有利于学生思辨水平和口头表达能力的提高，培养考生的逻辑思维和灵活应变能力。

### 三、改革评价体系

评分是考试的重要环节，对考试质量和功能发挥着重要作用。目前，我国学校普遍缺少一个综合评定人才的体系，考试分数只是直接反映学生在某一方面的学习情况，并不能真实全面地反映学生智能水平，考试成绩在某种程度上压制了学生的特殊才能和创造性。在考试改革中，要改变以考试成绩为标准的传统做法，评价体系必须对学生的知识、能力、素质作出全面的综合评价，考试成绩除采用绩点法或加权平均法外，还要加大对学生实践能力考核的比重，并对学生的创新、获奖、突出才能给予充分肯定，尊重学生个性发展，鼓励学生在各个方面“冒尖”、充分展示自身的聪明才干。

### 四、改革评价主体

在我国评价的主体一般是教师，而在国外对学生升学发挥作用的评价主体非常广泛。包括：学校校长、任课教师、家长和其他学生等，都可对学生的各科成绩、综合表现作出评价和推荐。因此，随着我高考制度改革的深入推进，随着社会条件和各种配套措施的落实，我们必须实行评价主体的多元化发展，从而有效促进评价的实施与效率。

### 五、树立诚信为本的社会道德

如美国的考试制度如此灵活、尤其是评价主体如此多元化竟然很少出现大规模的作假舞弊现象，令人惊讶。在中学树立这种

诚信为本的道德观，需要学校领导、教师甚至全社会的共同努力。

总之，我们在借鉴国外有关经验的基础上，还须结合我国国情，理性选择和慎重决策，坚持稳中求改，稳中求新。

# 第十五章　考试心理概述

考试是一种特殊的社会实践活动，对于考试的参与者——考生来说，考试过程实质上是一场心理活动，考试这一客观现实必然会在其头脑中产生各种各样的反映，这就是学生考试心理。具体地说，学生考试心理是指在考试这一特殊的情景下，考生在不停地联想已有知识与所考题目间的联系，进而寻找解决问题的方法这一过程中，所产生的复杂的心理活动，及由此引发的生理上的反应。各个学科的教学考试及一年一度的中考、高考，无论对于教学单位、教师还是学生来说，都具有重要的影响。如何让学生在考试中发挥正常的能力水平，不仅是考生所应思考的问题，作为中学教育工作者，除了做好教学工作外也应当积极借鉴考试心理学的理论，努力实践，培养学生形成良好的考试心理。

## 第一节　学生考试的心理状态

学生在考试的不同阶段——考试前、考试中和考试后，表现出不同的心理特点，不同心理状态的作用和表现形式各有不同。

### 一、考试前的心理状态

各种考试都会对考生的心理产生一定的影响，许多考生在考前很长一段时间就开始有了情绪反应，并且随着考试日期的接近，情绪波动幅度越来越大，一般会出现以下几类比较典型的心理

状态。

（一）积极的、乐观向上的心理状态

平时学习成绩较好、学习满意度较高，又得到教师和家长恰当鼓励的学生，其心理状态一般容易处于这个状态。这种学生面对考试，比较兴奋，但有节制，不紧张，乐于交谈有关考试的问题，对考试充满期待，对考试信心十足。有时还会有一些思维的小飞跃，例如出现观察力、记忆力提高，想象力丰富等好的心理反应。

（二）过于高估自己的心理状态

有些学生平时学习不踏实，对知识的学习浅尝辄止，而自己又盲目自信，过高地估计了自己的实力，超前沉浸在一种自我陶醉的成功感之中。往往越临近考试，他们越会感到忐忑不安，而一旦真正进入考场，就可能出现精神紧张、思绪混乱，平时能做出的题目，此时却可能一筹莫展。

（三）消极、敷衍的心理状态①

这种心理状态大多发生在平时学习成绩不太好的学生身上。处于这种心理状态的考生考前表现为情绪消沉、倦怠、浑身无力、反应迟钝、态度冷淡，对考试不感兴趣或厌倦，听到同学谈论考试问题会选择逃避，甚至产生放弃考试的想法。

针对考前情绪的紧张，尽量保持积极的、乐观向上的兴奋状态，对学生考试中水平的发挥有着重要的作用，学生可以尝试运用以下策略进行应对：①目的要正确，考前进行放松训练；②对自己的能力做恰如其分的估计；③减轻考生的精神压力；④注意活动、休息和营养；⑤制造轻松愉快的集体气氛。

① 彭剑飞．考试心理分析与指导［M］．长沙：湖南人民出版社，2001：91.

## 二、考试中的心理状态

考试过程是一个连续变化的过程，因受考试时间的限制，考试开始后，学生的心理一般要经历准备阶段、效率最高阶段和减力阶段。

### （一）准备阶段

从听到允许答卷的铃声开始，一般有几分钟时间。这是学生在考场上活动的启动阶段。在这一阶段中，考生从听到铃声响起后大脑即可进入兴奋状态，当接到试卷后，经过快速浏览，兴奋状态开始向以下几个方向转变。

1. 积极、自信的心理状态

当这种类型的考生粗略浏览试卷后，已经做到心中有数，发现其中只有很少一部分看上去不熟悉，便会平静下来，积极投身到试题中去，而且对自己充满信心，保持这种心理状态进行考试，考生的思维也会变活跃起来，对接下来的考试极为有利。

2. 消极、紧张的心理状态

有此种心理状态的考生，是因为当其快速浏览试卷后，发现许多试题难以解答，都是自己平时薄弱的地方，这时原先消极的心理暗示得到加强，于是便紧张起来，这将直接影响到下面的考试。

3. 漫不经心、无所谓的心理状态

处于这种心理状态的考生，一般都是学习困难生，他不会去浏览试题，也就不会提前紧张，只能按部就班进行答题，偶尔还东张西望、不认真考试。

### （二）效率最高阶段

这是学生的智力活动最积极的阶段，这个阶段延续时间的长短，对考试具有决定性意义。在这一阶段中，考生开始集中注意力，专心答题，此时的他们思维较活跃，不停地调动自己的已有

知识和平时积累的解题方法，智力活动最积极，也是效率最高阶段，但随着试题的交替变化，再加上调动知识过程的顺利与否，考生心理会发生多种变化。

（三）减力阶段

一般在试卷做完，或遇到解不出的难题以后，此时，情绪发生了新的变化，工作状况渐渐趋于不佳，在这一阶段，往往有两种对立的情绪表现，一是满意，二是失意。

考试中的效率最高阶段来得快，持续时间更长，势必会对考试越有利。我们可以采取以下的应对策略：①强迫自己冷静下来。②先易后难，步步为营。③每分必争，不浅尝辄止。

## 三、考试后的心理状态

前一科目考试结束以后，考生走出考场，一般都要迫不及待地和同学交换试卷答案，验证自己考试的结果，考得较理想的欢呼雀跃，考得不理想的可能会懊丧、气馁甚至妒嫉、愤恨。

这些不同的心理状态的产生、很大程度上决定于考生自认为的考试失败。一般说，考生会自认为前科考试成功会对以后的考试产生积极的，增力的作用。

如果前科考试失利以后，一般考生会产生消极的，减力的心理状态，有的考生甚至控制不住自己，一出考场就痛哭流涕，有的还企图放弃考试。

前科考试的成功也很可能引起盲目的乐观和过度的兴奋，导致以后考试的失利；而前科考试的失利，却可能会激励考生，使之产生积极的心理状态。

无论前科考试成功或失利，都要保持情绪的稳定，这就需要做到以下几点：

（一）继续努力、不骄不躁

在前科考试较理想的情况下，要努力克制自己骄傲的情绪，

不要过分得意。这时候应该多想一些困难，以抵消盲目的乐观，并借助于乐观的情绪，总结经验，积极准备下一科考试。

（二）不怕失利、乐观应对

一科考试的失利，在全部考试结束前只能算是暂时的和局部的，最后成败还未定论，只要认真总结教训，正确对付以后几门功课的考试，还有可能挽回损失。

## 第二节　影响考生考试的心理因素

学生在考试时表现出来的心理状态直接或间接影响到学生的考试行为及考试的结果。良好的考试心理对学生考试的发挥起到促进作用，不良的考试心理会对考试结果产生不良影响，考试行为和考试心理之间相互依存、相互影响、相互转换，并遵循着一定的规律，我们只有很好的认识这些规律，才能更好的认识和把握考试心理。

### 一、考试的心理动机

动机是需要刺激下直接推动人进行活动的内在心理动力。考生的动机如果不同，其应答态度、注意力、持久性和反映的速度就不相同，从而影响着考生的应试成绩。

考试动机的形成有很多原因。学生如果将考试与自己的名利、得失及命运联系在一起的时候，就会对考试寄予很高的期望，容易使学生产生积极向上的责任感和使命感，或产生一种消极的、影响考试正常发挥的制约作用①。因此，学生在考试过程中应该持有正确的态度，准确的评估自己，保持中等程度的期望值，从而

① 谷海军．学生的考试心理及应对策略［J］．陕西师范大学学报（哲学社会科学版），2005，S1：247－249.

保持适度的动机水平，对考生考试水平的发挥有着很重要的意义。

## 二、考生怯场心理

考试是一场紧张的心智活动，考生在考试时必须高度集中注意力，迅速地回忆和联想所学的有关知识，积极灵活地进行思维，有效地解决问题，才能取得较好的考试成绩。但怯场往往严重地影响考生正常水平的发挥，它是一种心理现象，是主客观条件变化对脑活动影响的结果。

### （一）怯场心理简述

怯场心理是考生因情绪过分兴奋或紧张而引起的无法自控的心理状态，考生在考试时因心理过度紧张等原因而出现心跳加快，血压升高，面红耳赤，头晕脑胀，思维迟钝，反常遗忘等生理、心理反应，在心理学上，怯场又称“考生竞技综合症”。

在重要的考试时，面对变化的考场环境、严肃紧张的考试气氛，或者考试中遇到了难题等，往往会出现不同程度的心跳加快、头胀耳鸣、面部发烧、口干冒汗、出现尿意，严重的则无法集中注意力，不能控制自己，表现为心慌意乱，手足无措，易发怒，或表现为信心不足、情感淡漠、动作呆板、头脑晕沉、思维迟缓等，连本来能答好的题也做不出了，以致不能充分发挥自己的原有水平。

### （二）引发怯场心理的因素

引发考生怯场心理的因素有很多，以下几种因素比较常见。

1. 过于看重考试

社会上普遍存在“望子成龙、望女成凤”的心理，家人对学生的期望值过高而不符合实际，学校普遍施行以升学为工作重心的教育评价机制，过分追求考试分数，考生自己过高的期望等都可能会使考生产生巨大的心理压力。

2. 不适应考试环境

有的考生的适应能力比较差，容易受到外界环境的干扰，一到陌生的考场便不适应，情绪特别紧张，总觉得有什么地方不舒服，连考场中的装饰和摆设都看不顺眼。

3. 紧张情绪互相影响

考生与同学之间的紧张情绪会相互影响、相互感染

4. 自我评价失实

一些考生在平常的考试中自我评价偏高，但因为实力不济，在中考或高考中遇到难题就乱了方寸，引发怯场。另外，一些考生自我评价偏低，缺乏自信，也容易引发怯场。

## 三、考生作弊心理

作弊是少数考生投机取巧心理的表现，是一种违纪行为。它是学生在某些外部原因影响下，通过一定的心理活动而形成的。作弊心理是作弊的知、情、意、行的发展所形成的。内外部环境的不协调可能会对学生的考试心理的发展形成不良影响，进而形成作弊心理。

(一) 作弊心理的主观原因

1. 中学生身心发展特征的影响

中学时期是学生身心发育的关键时期，中学生身心发育既有童年期的特征，又具有成年人的特征。到了高中阶段虽已具备成人的特征，但还未达到完全的成熟；在智力方面，中学生抽象逻辑思维能力迅速发展，并开始占主导地位，但具体形象成分仍起重要作用，思维的独立性和批判性显著发展起来，但认识能力还不高，仍带有片面性和表面性；因此在考试方面，中学生能够比较自觉地完成考试任务，但控制情感、自我监督的能力还不强。部分中学生由于过高地树立了自己的人生目标，没有充分认识到自身的实际水平和目标之间的距离，并且缺乏刻苦学习的精神，

在一次考试没有达到自己确立的目标后，不敢面对现实，进而产生虚荣心理，在下次考试中采取作弊的手段得到好成绩，以期望取得父母及老师的表扬和肯定。

2. 中学生的社会情感不成熟

社会情感是个体在社会生活中形成和发展起来的，包括道德感、理智感和美感。中学阶段是社会情感迅速发展的时期。中学生的社会情感虽然已经发展起来，但还不成熟。特别是由于中学生的道德认识水平的局限性，对道德标准的掌握不够全面和准确，容易产生错误的道德情感，缺乏正确的道德体验。在考试过程中，部分中学生对考试作弊的理解偏差，即产生偏差心理，认为考试作弊可以理解，必要时可以使用。

3. 中学生的厌学心理

厌学是指学生对学习的一种厌烦心理，它是一种比较持久而微弱的学习生活的情绪状态。具有厌学心理的中学生表现为对学习不感兴趣，听不进课，课后不能认真完成作业，讨厌考试，将考试视为自己的对立面，以致于“谈考色变”。但由于受到家庭等多方面制约，不得不去参加考试，于是便产生了强迫心理，不管考什么科目，满脑子想到就是找机会作弊。

4. 中学生的情绪两极性

中学生的情绪情感表现比较热烈，带有明显的两极性特征，时常因为一些小事被感动得热泪盈眶，或者是被激怒而情绪暴躁。中学生对同学关系相当重视，富有同情心。学习较为优秀的学生看到有些同学老是不及格，受到家长和老师的斥责甚至体罚，于心不忍。因而在考试过程中，如果这类同学请求帮忙时，他们便产生同情心，进而为同学提供作弊材料，帮助其作弊。

（二）作弊心理的客观原因

1. 学校因素的影响

学校是学生成长的重要环境之一，是学生进行学习的主要场

所。导致中学生考试作弊心理的因素有很多种，学校教育工作的失误是其中重要的一种。具体来说，表现在以下几个方面。

①学校片面追求升学率

随着新课程改革的不断深入开展，素质教育的实施在我国取得了很大的成绩，但受多年施行的应试教育的影响，在教学工作中仍以中考、高考为风向标，以学生的考试成绩作为对学生及教师的评价标准，过分的强化了考试分数的作用。这使得学生和教师一味地追求考试分数，从而产生不良的影响。

②学校考试疏于管理

有些学校对考试管理工作重视程度不够，甚至相关的管理制度不健全，缺乏有效的监督体制；有些学校的监考教师的责任心不强，对作弊学生放任自流，进而造成考试作弊学生放纵妄为，形成恶性循环；有些学校忽视考试的考前教育，忽视考试纪律教育和诚信教育，造成部分中学生认不清考试作弊的严重性和危害性，进而产生侥幸、从众的作弊心理。

2. 家庭因素的影响

在当前社会竞争压力越来越大的影响下，几乎所有的家长对子女寄予了过高期望，而家长的期望水平不仅影响着子女的学业成就，也同时影响着中学生心理及道德的形成。过高的期望值往往是不现实的，可能会对学生产生较大的心理压力，致使子女为了达到家长的期望分数，产生考试作弊心理，实施作弊行为。

3. 社会风气和环境的影响

全社会教育思想没有端正，即教育没有真正走到素质教育的轨道上来，是作弊心理赖以滋生的主要社会背景。社会各种不正常考试考评现象是其产生的参照物。一个时期以来，全国各行各业的各种考试蜂拥而起，一股强大的作弊浊流亦泛滥成灾，对学生来说这些各级各类不正常的成人考试考评制度暗示了考试作弊的实用性，在群体传染下，学生考试作弊心理滋生有了其可行性。

## 四、考试焦虑心理

考试焦虑是在一定的应试情境激发下，受个体评价能力、人格特征与其它身心因素所制约，以担忧为基本特征，以防御或逃避为行为方式，通过不同的情绪反应所表现出来的一种心理状态。考试焦虑，不管是正常焦虑还是神经过敏性焦虑，都是对个体自尊心的一种威胁，它会激起个体改变自身目前的唤醒状态，从而去努力达到某个目标。因此，考试焦虑本身并不是一种病态的问题，有时候焦虑是一种正常的情绪反应，适当的考试焦虑还有利于发挥潜能、解决问题，成为应考有效学习的动力之一。但其反应在强度上过高或持续的时间过久，就会影响考生在考试期间正常的学习与生活，或继续引发身心上的焦虑症状，使得个体百般困扰而需要接受治疗[①]。

### （一）考试焦虑的级别

考试焦虑可分为四级。

第一级：镇定状态。考生头脑清醒、思想重视、态度端正、积极乐观、坦然自如，但有时略有激动，具有紧迫感。

第二级：轻度焦虑。考生心情较为激动，有一种焦虑感，心中有点惶恐不安。这种焦虑有功无过，只要考生把它控制在比较短的时间内，是不会影响其心理健康的。

第三级：中度焦虑。心情过于激动，焦虑过高。说明考生的神经开始有些紊乱，身心健康已经有所损害。

第四级：重度焦虑。考生对考试已产生恐惧心理，焦虑感很强，已患上“考试焦虑症”。

### （二）考试焦虑产生的原因

在当前新课程改革的背景下，各种升学考试激烈程度仍在增

① 宋兴川，李丽虹．学会自我调节克服考试心理困惑［J］．湖北招生考试，2004，（8）：13－18.

加，学生的焦虑心理出现了上升的趋势，部分考生在考试前或考试中常产生忧虑、烦躁、紧张甚至恐惧，出现思维受阻、身体不舒服、失眠、食欲不振状况等。

考试焦虑既可能来源于考生内在的压力，如考生给自己制定的成就目标过高，害怕失败；也可能是考生对自我的认识不足，缺乏自信心，过分在意他人对自己的评价；也可能来自于外界的压力，如家长过高的期望值，造成学生在考前出现忧心忡忡的不安情绪；也可能来源于教师教育及教学的方式，当教师对学生寄予了一定的厚望，而学生却“试”与愿违，来自教师的这种外力极易使学生在考试中产生焦虑情绪。另外，考试中强烈的竞争力，也常常使学生被分数所制约，出现分数至上的状况。当分数成为测试学生能力的惟一指标时，迫于压力，学生也会产生一定的焦虑情绪。

（三）考试焦虑的自测

了解了考试焦虑的定义及产生的原因，学生应该进一步了解自己当前的状态，下面的测验旨在对中学生的考试焦虑心理做客观的诊断。

考试焦虑程度自我测查表

【指导语】下面共有33道题目，每题有4个备选答案，请根据自己的实际情况，在题目后面圈出相应字母，每题只能选择一个答案，其相应字母的意义是：

A. 很符合自己的情况。　　B. 比较符合自己的情况

C. 较不符合自己的情况　　D. 很不符合自己的情况

1. 在重要的考试前几天，我就坐立不安了。　　A B C D
2. 临近考试时，我就泻肚子了。　　A B C D
3. 一想到考试即将来临，身体就会发僵。　　A B C D

4. 在考试前，我总感到苦恼。 A B C D

5. 在考试前，我感到烦躁，脾气变坏。 A B C D

6. 在紧张的温课期间，常会想到：这次考试要是得到个坏分数怎么办? A B C D

7. 越临近考试，我的注意力越难集中。 A B C D

8. 一想到马上就要考试了，参加任何文娱活动都感到没劲。 A B C D

9. 在考试前，我总预感到这次考试将要考坏。 A B C D

10. 在考试前，我常做关于考试的梦。 A B C D

11. 到了考试那天，我就不安起来。 A B C D

12. 当听到开始考试的铃声响了，我的心马上紧张地急跳起来。 A B C D

13. 遇到重要的考试，我的脑子就变得比平时迟钝。 A B C D

14. 看到考试题目越多、越难，我越感到不安。 A B C D

15. 在考试中，我的手会变得冰凉。 A B C D

16. 在考试时，我感到十分紧张。 A B C D

17. 一遇到很难的考试，我就担心自己会不及格。 A B C D

18. 在紧张的考试中，我却会想些与考试无关的事情，不能集中注意力。 A B C D

19. 在考试时，我会紧张得连平时记得滚瓜烂熟的知识一点也想不起来。 A B C D

20. 在考试中，我会沉浸在空想之中，一时忘了自己是在考试。 A B C D

21. 考试中，我想上厕所的次数比平时多些。 A B C D

22. 考试时，即使不热，我也会浑身出汗。 A B C D

23. 在考试时，我紧张得手发僵，写字不流畅。 A B C D

24. 考试时，我经常会看错题目。 A B C D

25. 在进行重要的考试时，我的头就会痛起来。 A B C D

26. 发现剩下的时间来不及做完全部考题，我就急得手足无措。 A B C D

27. 如果我考了个坏分数，家长或教师会严厉地指责我。 A B C D

28. 在考试后，发现自己懂得的题没有答对时，就十分生自己的气。 A B C D

29. 有几次在重要的考试之后，我腹泻了。 A B C D

30. 我对考试十分厌烦。 A B C D

31. 只要考试不记成绩，我就会喜欢进行考试。 A B C D

32. 考试不应当像现在这样的紧张状态下进行。 A B C D

33. 不进行考试，我能学到更多的知识。 A B C D

计分与评价：

统计你所圈各个字母的次数，每圈一个 A 得 3 分、B 得 2 分、C 得 1 分、D 得 0 分。用下列公式可以算出你的总得分：

总得分 = 3 × 圈 A 的次数 + 2 × 圈 B 的次数 + 圈 C 的次数

根据你的总得分查下面的评价表，就可以知道你的考试焦虑水平。

**表 15－1　评价表**

| 总分 | 焦虑水平 |
| --- | --- |
| 0 ~ 24 | 镇　　定 |
| 25 ~ 49 | 轻度焦虑 |
| 50 ~ 74 | 中度焦虑 |
| 75 ~ 99 | 重度焦虑 |

如果学生通过利用前面的焦虑程度自查表，判断自己确实有考试焦虑，下面的这个量表可以帮助他们分析考试焦虑的原因。

## 考试焦虑自测量表①

【指导语】为了帮助你准确地把握自己在考试焦虑方面存在的问题，在此准备了这份考试焦虑自我检查表。请你仔细阅读每一道题目，看看它是否反映出你在应试时的经验和感受。如果是的话，请在该题目右边上画一个“√”；如果不是的话，则无需做任何标记；如果有些题目实在难以确定，请你画一个“?”。请如实作答，不要花太长时间思考，尽可能回答你看完题目后的第一印象。

1. 我希望不用参加考试就能取得成功。

2. 在某一考试中取得的好分数，似乎不能增加我在其他考试中的自信心。

3. 人们（家长、老师和朋友等）都期待我在考试中取得成功。

4. 考试期间，有时我会产生许多对答题毫无帮助的莫名其妙的想法。

5. 重大考试前后我吃不下东西。

6. 对喜欢向学生搞“突然袭击式”考试的教师，我始终感到害怕。

7. 在我看来，考试过程似乎不应该搞得太正规，因为那样容易使人紧张。

8. 一般来说，考试好的人将来必定在社会上取得较好的地位。

9. 重大考试之前或者考试期间，我常常会想到其他人比自己强多了。

10. 如果我考糟了，即使自己不会老是记挂着它，也会担心别人对自己的评价。

---

① 耿文秀．决胜考场——考试心理自助［M］．上海：华东师范大学出版社，2001：36－40.

11. 对考试结果的担忧，在考试前妨碍我准备功课，在考试中妨碍我答题。

12. 面临一场必须参加的重大考试，我会紧张得睡不好觉。

13. 考试的时候，如果监考老师来回走动并注视着我，我便无法好好集中思想答题。

14. 如果考试被废除，我想我的功课实际上会学得更好。

15. 当了解到考试结果的好坏将在一定程度上影响我的前途时，我就会心烦意乱。

16. 我知道，如果自己能集中精神，考试时便能超过大多数人。

17. 如果我考得不好，人们将对我的能力产生怀疑。

18. 我似乎从来没有为参加考试进行过充分的复习准备。

19. 考试前，我的身体不能放松。

20. 越是重大的考试，我的大脑越像一锅粥。

21. 考场中的噪音（比如日光灯的电流声、送暖气或冷气的声音、前排应试者发出的声音等等），会令我大为烦恼。

22. 考试前，我会有一种空虚、不安的感觉。

23. 考试使我对能否达到自己的目标产生怀疑。

24. 考试其实并不能反映一个人对知识的掌握状况和程度。

25. 如果考试得了低分数，我不愿把自己的确切分数告诉任何人。

26. 考试前，我常常感到还需要再充实一些知识。

27. 重大考试之前，我的胃会不舒服。

28. 有时，在参加一次重要考试的时候，一有某些消极的想法，我似乎整个人都会垮了。

29. 在即将得知考试结果前．我会感到十分焦虑或不安。

30. 但愿将来我能找到一个不需要考试便能被录用的工作。

31. 假如在这次考试中我考得并不好，我想这意味着自己并不

像原来所想像的那样聪明。

32. 如果我的考试分数低，我的父亲或母亲将会感到非常失望。

33. 对考试的焦虑简直使我不想认真准备了，这种想法又使我更加焦虑。

34. 应试时我常常发现自己的手指在哆嗦，或者双脚在打颤。

35. 考试过后，我常常感到本来自己应该可以考得更好一些。

36. 考试时，我情绪紧张，妨碍了注意力的集中。

37. 在某些考试题上我费劲越多，脑子越乱。

38. 如果我考糟了，且不说别人会对我有看法，就是我自己也会对自己失去信心的。

39. 应试的时候，我身体某些部位的肌肉很紧张。

40. 考试之前，我感到缺乏信心，精神紧张。

41. 如果我的考试分数低，我的朋友会对我感到失望。

42. 在考前，我所存在的问题之一是不能确知自己是否作好了准备。

43. 当我必须参加一次确实很重要的考试时，我常常感到恐慌。

44. 我希望主考人能够察觉，参加考试的某些人比另外—些人更紧张，我还希望主考人在评价考试结果时能将上述因素考虑在内。

45. 我宁愿写一篇论文，也不愿参加考试。

46. 在公布我的考分前，我很想知道别人考得怎么样。

47. 如果我得了低分数，我认识的某些人将会感到快活，这使我心烦意乱。

48. 我想，如果我能单独进行考试，或者没有时限压力的话．我的成绩就会好得多。

49. 考试成绩直接关系到我的前途和命运。

50. 考试期间，有时我非常紧张，以至于忘记了自己本来知道的知识。

请对照表15－2的标准答案，从中你可以发现引起自己考试焦虑的最重要的原因及具体的焦虑反应。

**表15－2　考试焦虑自我检查表的标准答案**

| 类别 | 测查的内容 | 相关的题目序号 |
| --- | --- | --- |
| 考试焦虑的来源 | 担心考砸后他人对自己的评价 | 3、10、17、25、32、41、46、47 |
| | 担心失败对个人的自我印象增加威胁，降低自我评价 | 2、9、16、24、31、38、40 |
| | 担心未来的前途 | 1、8、15、23、30、49 |
| | 担心对应试准备不足 | 6、11、18、26、33、42 |
| 考试焦虑的表现 | 身体反应 | 5、12、19、27、34、39、43 |
| | 思维阻抑 | 4、13、20、21、28、35、36、37、48、50 |
| 其他 | 一般性的考试焦虑 | 7、14、22、29、44、45 |

## 第三节　考生不良考试心理的预防与辅导

一个成功的应试者除了应具备一定的智能水平、知识水平及良好的生理状况外，还应当具有稳定的心理素质，心理素质的稳定与否是影响考试活动顺利进行和良好考试成绩取得的重要因素。在当前社会和学校的环境中，中学生依然面临着强大的升学考试压力，普遍存在着一定的考试心理问题，如何预防和治疗学生的不良考试心理，也是教育工作者需要认真研究的问题。

### 一、考生怯场心理的应对策略

怯场现象是一种常见的考试心理现象，是主客观条件变化对

脑活动影响的结果。怯场心理对考生应试水平的发挥有一定的影响，如何减少其影响，我们可以从以下几个方面提出相应解决对策。

(一) 有一个正确的学习和应试动机

动机的状态直接影响着解决问题的效率，我们可以绘成一条倒转的U型曲线，即问题解决者动机增加，解决问题的效率也随着增加，直到达到一个最高点。超过了这一点，动机强度的增加，却会造成解决问题的能力降低。可见要收到考试的效果，就要有合适程度的动机，动机是推动人们去行动的内在力量。过弱，对考试敷衍了事，当然考不好；过强，把考试视为决定自己命运前途的关键，从而背上沉重的思想包袱，考期愈近愈紧张，考试中一旦遇到不顺利，就容易产生焦躁情绪，紧张不安，在大脑产生一个影响思维的优势兴奋灶，从而发生怯场现象。

(二) 系统地牢固地掌握基本知识和技能

由于平时学习与训练，知识系统已在大脑皮层形成了较为巩固的暂时神经联系系统，如在迎考复习中又得到了很好的复习，这个知识系统就更加巩固，就不易被干扰、被抑制。

(三) 积累考场经验，合理调控自己的情绪变化

考试的条件、环境变化也能引起怯场的刺激。例如，全新的考场、陌生的监考人，甚至试卷的纸张，印刷的形式等方面的变化，都能在应试者大脑皮层引起干扰正常思维的优势兴奋灶。面对这些客观条件的变化，就要求学校、教师帮助学生创造条件以丰富考场经验。可给学生讲述临场的心理状态，通常还可以进行模拟正式考试的情境。应试者在平时测验、考试时要认真对待，想象重大考试的情境，即使更换了环境、条件也一样从容不迫。

(四) 采取注意中心转移法，使暂时遗忘解除

由于考场的气氛或是试题难度的刺激，在大脑皮层引起的不利于思想的兴奋灶，会抑制原有的知识，发生暂时的遗忘。对付

这种现象的方法，可以采用使兴奋转移的方法。

（五）保持正常的生活规律，避免人为的紧张

临考试的准备工作和生活规律与怯场现象有关。人的生活是有一定规律的，在正常生活中获得的知识，在正常环境中才能得以更好的发挥和使用。

## 二、考试作弊心理的预防策略

中学生考试作弊心理的预防是一项繁杂的系统工程，涉及方面广，社会影响大，如不能及时有效的预防和矫治，极易对中学生的个体成长产生负面的影响。考试作弊问题不可能避免，但这并不意味着面对考试作弊我们可以放任自流，只要措施得力，方法正确，我们还是能够减少中学生考试作弊现象的出现。从中学生考试作弊心理的形成原因看，必须从学校、家庭、社会三个方面做好系统预防工作。

（一）学校工作方面

1. 加强德育教育工作

学校应加强对中学生进行理想、动机、人生观和价值观等方面的教育，增强中学生的责任感和使命感，帮助中学生明确学习目标，端正学习动机，培养坚强的学习意志。同时，要提高中学生对考试作弊的严重性、危害性的认识。要让其认识到考试作弊是一种不道德的行为，帮助他们树立良好的道德意识，加强自身修养。要加强正面教育，提倡从我做起，自觉抵制考试作弊。

2. 加强学生意志训练

教师要使学生了解意志在成长过程中的作用，有意识地进行意志训练。一方面，要善于引导中学生抵制不符合行动目的的主客观诱因的干扰，另一方面要善于激发中学生长久地维持自己开始的符合目的的行动。要努力培养中学生具有自信、自治、果断、持久性和勇于克服困难的意志品质，注重引导学生克服盲目从众、

言行不一等不良意志品质。

3. 缓解应试心理压力

研究表明适度的心理压力是学生进步的内驱力，是促进学生主动发展的动力。但如果教师和家长给中学生的学习压力太大，持续时间过长，将会使压力变成阻力，从而不利于中学生的学习和身心健康。为缓解中学生的心理压力，教师要教育学生正确认识自己，学会学习，能给自己设立一个适当的努力目标，从学生本身减小内在的心理压力。教师要掌握学生的思想动向，进行必要的考前诚信教育，对重点学生进行重点思想工作，通过精神鼓励和感化，解决他们存在的思想问题，调整其情绪和心理状态，消除精神负担，认识到自己在考场上的责任，提高自尊心、自信心和自制力。

(二) 家庭工作方面

1. 遵循教育原则

在对子女的教育中，科学的教育方法和管教方式是十分重要的，应该采取平等、民主和宽严适度的教育方法，而不是专横、独断、惟命是从的教育方法，更不能过分放纵、溺爱、百依百顺。中学生考试作弊的重要因素，就是家长的教育方法不恰当。

2. 树立良好榜样

作为孩子第一位老师的家长，是子女仿效的最直接的对象。家长必须要以身作则，树立好的榜样。否则，家长素质低下，言谈举止不文明，或在道德品质、行为习惯上有问题，会对子女造成不良影响。

## 三、考试焦虑的心理辅导

由于考试焦虑是一种由多种因素造成的心理失调现象，它不仅与学校、家庭及社会有关，而且还与学生个体自身的身体素质和心理素质等有很大的关系，所以要降低学生的考试焦虑水平，

必须内外结合，多方努力才能够显现效果。

（一）考试焦虑的认知矫正

针对学生的考试焦虑，认知矫正是一种重要的辅助途径。其主要目的是通过影响学生对考试的知觉和评价，使学生对考试采取一种现实合理的态度，进而对“即将面对考试”这一情景产生一种控制感。大量的研究表明：一个人对即将面对的生活事件的可控制预期会有效减轻焦虑的反应。教师要指导学生进行认知矫正，帮助他们发现自己的长处，肯定自己的优点，增强学生的学习信心，进而对考试情景产生一种控制感，减轻考试焦虑。

针对学生的考试焦虑心理，可以按照以下四步骤自助性认知矫正程序进行矫正：

①引导学生检查自己对考试的担忧，让学生把自己想起的一切担忧全部写下来。然后再整理写出的“担忧”，去掉重复的内容，把实质相同的合并起来。再让学生按“担忧”尺度的大小依次排列各项目。

②引导学生对自己的担忧进行合理性分析。如自己的担忧是否有事实依据？自己的担忧是否以偏概全？自己的担忧是否极端化？

③引导学生与“担忧”质辩。要求学生针对担忧的不合理之处，以事实、理性的分析予以辩驳。

④得出合理的反应。以上述三个步骤为基础，帮助学生合理地认知“担忧”，并树立积极的态度。

（二）考试焦虑的行为矫正

1. 区别情况对症下药

（1）重度焦虑学生

教师要针对学生个性特点和认知水平，制定一份针对性强的训练计划，进行重点辅导，在加倍的关爱下，使他们冷静地面对焦虑，逐步提高对心理挫折的耐受力和对焦虑的适应能力，激发

克服焦虑的勇气。

（2）中度焦虑学生

教师要善于观察、善于发现每个学生产生焦虑的原因，做到对症下药。对于那些平时情绪较稳定的学生，如果突然出现过度焦虑的某种表现，要及时查清原因，对他们要特别的关注爱护，实施调整措施。对由于学习成绩差，遭到老师和同学严重歧视的原因造成的，要采取适当的补救措施，多鼓励、少批评，从而消除其心理负担。

（3）轻度焦虑学生

高考等大型的考试前出现轻度焦虑也在所难免，其焦虑水平在正常范围。

2. 帮助树立自信心

自信的学生比焦虑的学生更容易取得较好的成绩，帮助学生树立自信心对缓解考试焦虑有着重要的意义。

（1）学会正确评价自己

要对自己各方面的情况有客观准确的估计，养成自我监督情感的习惯，提高调节和控制情感的能力，通过自我正确评价，确立适当的目标，但不因目标过高而沮丧，在心理压力过大、烦躁不安时，学会表达和宣泄，这样间接地就缓解了过度的焦虑。

（2）考试动机适当

以平常的心对待高考，从心理上消除对考试的不必要顾虑，期望值不要过高，不要过多考虑一定要考到多少分，考上什么院校等。

3. 冷静、勇敢地面对焦虑

制定平时、考前、考中的复习目标和规划来代替焦虑，做到每个阶段能自我复习、自我检测。高考的特点是“重基础，考能力”，因此，最好能按各科教学大纲与考试大纲的要求进行全面系统的复习，平时系统消化，考前综合归纳，决不能凭侥幸心理靠

押题取胜。

4. 掌握考试技巧，从容应付考试

善于总结平时临场考试时的审题答题顺序、技巧、书写表述、检验等各环节的成与败，选择适合自身特点的心理调节方式，把平时当高考，把高考当平时，避免紧张与焦虑。

5. 预防不良的外来“刺激”

不良的外来“刺激”会给学生心理带来干扰，使其惊慌失措，心理紧张，影响正常思维，使学生原来已知信息难于提取，引发焦虑反应。因此要预防不良的外来“刺激”，减小心理干扰，减少产生考试焦虑心理的情境。

（三）考试焦虑的预防策略

在考试期间，焦虑的存在与出现是必然的，这是因为中等程度的焦虑是激励个体努力复习的动力，如何将其维持在合理的程度与范围内，不致妨碍个人复习应考的心理状态，是每位考生所必须了解与实践的任务。这里从预防的角度提供一些日常生活中可行的注意措施：①重视身体的检查，通过检查，可及早发现导致紧张不安、烦躁的身心原因，因势利导预防焦虑的进一步发展。②心情郁闷焦躁不安时，不妨与父母、老师、知心朋友交谈倾诉，可以减缓内在的压力。学习放松的技巧，养成放松的习惯，在需要时，短暂的放松训练会产生意想不到的效果。③精力充沛的身体是维持日常生活的基础，补充必要的养分，才有精神专注于工作；水果和牛奶都有除烦安神的作用。④养成每天运动的习惯，任何运动项目皆有助于身心发展，减缓紧张，消除焦虑。⑤睡得好与吃得好同样重要，适宜的睡眠可松弛紧张的情绪，缓和承受的压力，储备面对挑战的精力。⑥有效安排自己的学习计划与作息时间，可使精力得到充分的发挥，减少紧张与焦虑的累积。⑦考虑自己的能力与条件，制定切实可行的学习目标与成绩比较标准。

（四）克服考前焦虑的训练

考生在临考前出现焦虑，是一种很正常的现象，关键在于焦虑的强度。高度的焦虑可以使考生长期处于一种极度的紧张状态，产生对考试的恐惧；而低度的焦虑则会使考生对即将进行的考试无所谓，引不起应有的重视。心理学研究表明：中等程度的焦虑对提高考试效果最为有效。考试作为一项复杂的脑力劳动，需要考生保持清醒的头脑和中等程度的焦虑，而强烈的求胜动机只能引起考生较强的情绪兴奋和高度焦虑。因此，要保持临考前中等强度的焦虑，必须努力做到：

1. 降低求胜动机

只有适当降低求胜动机，才能提高理智水平，消除焦虑带来的负面作用。一般情况下，可以采用焦点转移法，即找到一种或一类可以替代注意焦点的第二目标，以减轻不达目的不罢休的心理压力。例如对待升学考试，可以联想一些有志者受挫折而另辟新途取得成功的事例，真正认识到上大学虽然较为理想，但却不是唯一的选择，从而降低注意焦点的诱惑作用。当然也不宜将追求目标说得一无是处，否则，就会失去考试成功的心理动力。

2. 学会自我放松

考试是场艰苦的劳动，需要消耗大量的精力。因此，临考前应减少运动量和情绪波动，努力使自己的身体处于一种宁静、和谐的静息状态。研究证明，在这种静息状态下，利于全面回忆学过的知识结构，找出遗忘点，进而对症复习，这比体育锻炼、完全休息或临阵磨枪都更有利于提高考试效果。

3. 作好充分准备

考试是一种最容易发生情绪波动的活动，任何一点小小的刺激都可能在考生的心理上激起轩然大波。因此，考生必须重视考前的各种准备工作。

（1）复习准备

考生应有计划地开展复习任务，精确地规定出每一阶段的复习进度，可以把在一小时或半小时里，打算完成的任务规定下来，并规定出完成复习任务的最后期限。这样，就不至于在心理上感到复习任务是那么艰巨而可怕了。

（2）物质准备

如周密地准备好考场答题时必需的各种文具，事先熟悉考场环境，以及充分估计在考场上可能发生的生理、心理变化异常或意外事件，如出汗、头晕、呕吐、汗渍、墨渍等，事先作好应变措施。

（3）精神准备

除了前述的减轻心理负担的方法外，还要对于考试时可能出现的猜题有误、记忆卡壳、以及如何去解答各类试题等都要有所准备，制定对策。如注意以整洁工整的试卷给阅卷者以较好的第一印象；注意解题采用先易后难的顺序；注意处理好求准、求快的关系，力争一遍成功；注意答案简明扼要，谨防画蛇添足；注意合理安排时间，细心复查，不提前交卷；注意对于无力解答的题，只要不倒扣分，即使猜测回答，也不留空白等。

经过上述的心理训练与周密准备，有利于考生解除疲劳、消除紧张、从而精力充沛、信心十足地去参加考试。

# 主要参考文献


[1] 沈兼士. 中国考试制度史（第四版）[M]. 台湾：台湾商务印书馆，1980.

[2] 叶沛华，陈一百等. 教育统计学 [M]. 北京：人民教育出版社，1982.

[3]（英）罗伯特·蒙哥马利（Robet Montgomery）. 黄鸣译，考试的新探索 [M]. 南宁：广西人民出版社，1984.

[4] 中国大百科全书总编辑委员会（教育编辑委员会）. 中国大百科全书（教育）[M]. 上海：中国大百科全书出版社，1985.

[5] 张植信，赵发，王利生等. 百科知识竞赛大全 [M]. 太原：山西科学技术出版社，1986.

[6] 于信凤. 考试学引论 [M]. 沈阳：辽宁人民出版社，1987.

[7]（美）B. S 布卢姆著，张云高译. 对学习者的新看法 [M]. 上海：华东师范大学出版社，1987.

[8] 盛奇秀. 中国古代考试制度 [M]. 济南：山东教育出版社，1988.

[9] 郑日昌. 中学生心理诊断 [M]. 济南：山东教育出版社，1988.

[10] 金玉章. 实用教育统计与测量 [M]. 北京：教育科学出版社，1988.

[11] 徐玖平．考试学［M］．成都：成都科技大学出版社，1989.

[12] 瞿葆奎．教育学文集（教育评价卷）［M］．北京：人民教育出版社，1989.

[13]（美）布卢姆著，罗黎辉译．教育目标分类学（认知领域）［M］．上海：华东师范大学出版社，1989.

[14] 顾明远．教育大辞典（第一卷）［M］．上海：上海教育出版社，1990.

[15] 金诤．科举制度与中国文化［M］．上海：上海人民出版社，1990.

[16] 张敏强，刘昕．标准化考试［M］．北京：高等教育出版社，1990.

[17] 陈友新，王建新．考试心理学［M］．哈尔滨：黑龙江科学技术出版社，1990.

[18] 安文铸．学校管理辞典［M］．北京：中国科学技术出版社，1991.

[19] 刘肃毅．面试教程［M］．北京：北京农业大学出版社，1992.

[20] 于钦波，刘民．考试学概论［M］．沈阳：辽宁教育出版社，1992.

[21] 张厚粲，刘昕．考试改革与标准参照测验［M］．沈阳：辽宁教育出版社，1992.

[22] 杨学为，朱仇美，张海鹏．中国考试制度史资料选编［M］．合肥：黄山书社，1992.

[23] 辞海编辑委员会．辞海（缩印本）［M］．上海：上海辞书出版社，1993.

[24] 张希清．中国科举考试制度［M］．北京：新华出版社，1993.

[25] 罗增儒．数学竞赛教程 [M]．西安：陕西师范大学出版社，1993.

[26] 王孝玲．教育统计学（修订本）[M]．上海：华东师范大学出版社，1993.

[27] 刘存侠，高安民．教育统计与测量 [M]．西安：陕西师范大学出版社，1993.

[28] 张厚粲．心理与教育统计学 [M]．北京：北京师范大学出版社，1993.

[29] 顾援，贺斌．考试心理与应考技巧 [M]．太原：山西高校联合出版社，1993.

[30] 贺贤孝，杨树林．数学考试论 [M]．南宁：广西教育出版社，1994.

[31] 孙瑞清，胡大同．奥林匹克数学教学概论 [M]．北京：北京大学出版社，1994.

[32] 杨成鉴，金涛声．中国考试学 [M]．北京：书目文献出版社，1995.

[33] 谢青．中国考试制度史 [M]．合肥：黄山书社，1995.

[34] 刘芃．历史学科考试测量的理论与实践 [M]．北京：人民教育出版社，1996.

[35] 刘海峰．科举考试的教育视角 [M]．武汉：湖北教育出版社，1996.

[36] 韦志成．语文考试论 [M]．南宁：广西教育出版社，1996.

[37] 薛理银．教育信息处理基础 [M]．北京：北京师范大学出版社，1996.

[38] 陆震．中外学校教育考试制度探讨 [M]．北京：高等教育出版社，1997.

[39] 郭齐家．中国古代考试制度 [M]．北京：商务印书馆，

1997.

[40] 邵瑞珍．教育心理学 [M]．上海：上海教育出版社，1997.

[41] 周宏立．化学竞赛指导与人才的发现 [M]．哈尔滨：黑龙江教育出版社，1997.

[42] 柏授．考试心理 [M]．北京：中国人口出版社，1997.

[43] 张厚粲．教育测量学 [M]．北京：人民教育出版社，1998.

[44] 郭湛．人活动的效率 [M]．北京：人民出版社，1999.

[45] 徐升．人才测评 [M]．北京：企业管理出版社，2000.

[46] 卢纹岱．统计分析 [M]．北京：电子工业出版社，2000.

[47] 杨学为．中国考试改革研究 [M]．北京：北京大学出版社，2001.

[48] 耿文秀．决胜考场——考试心理自助 [M]．上海：华东师范大学出版社，2001.

[49] 彭剑飞．考试心理分析与指导 [M]．长沙：湖南人民出版社，2001.

[50] 陈玉琨．课程改革与课程评价 [M]．北京：教育科学出版社，2001.

[51] 刘海峰．中国考试发展史 [M]．武汉：华中师范大学出版社，2002.

[52] 周卫勇．走向发展性课程评价 [M]．北京：北京大学出版社，2002.

[53] 黄光扬．教育测量与评价 [M]．上海：华东师范大学出版社，2002.

[54] 康乃美，蔡炽昌等．中外考试制度比较研究 [M]．武汉：华中师范大学出版社，2002.

[55] 贺祖斌. 现代教育统计测量与评价 [M]. 桂林：广西师范大学出版社，2001.

[56] 金娣，王钢. 教育评价与测量 [M]. 北京：教育科学出版社，2002.

[57] 凌云. 考试统计学 [M]. 武汉：华中师范大学出版社，2002.

[58] 谢国栋. 突破考试心理关 [M]. 上海：百家出版社，2002.

[59] 漆书青. 现代测量理论在考试中的应用 [M]. 武汉：华中师范大学出版社，2003.

[60] 廖平胜. 考试学原理 [M]. 武汉：华中师范大学出版社，2003.

[61] 廖平胜. 考试是一门科学 [M]. 武汉：华中师范大学出版社，2003.

[62] 李安，苏建祥. 高中化学决赛辅导 [M]. 西安：陕西师范大学出版社，2003.

[63] 申燕萍. 考试心理训练 60 法 [M]. 上海：上海大学出版社，2003.

[64] 海玲. 笑傲考场 [M]. 上海：东华大学出版社，2004.

[65] 闫蒙钢. 化学教学测量与评价导论 [M]. 北京：北京科学技术出版社，2004.

[66] 皮连生. 教育心理学 [M]. 上海：上海教育出版社，2004.

[67] (美) B. S 布鲁姆等著，邱渊等译. 教育评价 [M]. 北京：北京师范大学出版社，2004.

[68] 潘永庆等. 多元评价：创新教育的有效机制 [M]. 济南：山东教育出版社，2004.

[69] 高凌飚. 普通高中新课程——模块学业评价 [M]. 北

京：高等教育出版社，2005.

[70] 闫蒙钢. 中学化学课程改革概论 [M]. 合肥：安徽人民出版社，2006.

[71] 谢小庆，张晋军. 考试研究文集（第三辑） [M]. 北京：经济科学出版社，2006.

[72] 全国中学物理竞赛委员会办公室. 全国中学生物理竞赛专辑（1997－2004）[M]. 北京：教育出版社，2006.

[73] 上海市教育考试院. 招生考试研究 [M]. 上海：上海教育出版社，2007.

[74] 鄢明明. 大规模考试的演变与育人——论会考与高考的改革 [D]. 武汉：华中师范大学，2003.

[75] 江琦. 中学生考试心理问题及教育对策研究 [D]. 重庆：西南师范大学，2003.

[76] 李艳平. 中学生考试焦虑与心理健康、学业成绩的相关研究 [D]. 上海：上海师范大学，2003.

[77] 孙云鹏. 中学生考试作弊心理的成因及预防 [D]. 石家庄：河北师范大学，2005.

[78] 胡向东. 民国时期中国考试制度的转型与构建 [D]. 武汉：华中师范大学，2006.

[79] 冯华. 高中招生制度改革势在必行——访教育部初中毕业与高中招生制度改革项目组协调人贺斌研究员 [N]. 中国教育报，2004－06－09.

[80] 钟启泉. 国外的“教育目标”说及其启示 [J]. 外国教育，1980，(3)：19－21.

[81] 廖平胜. 论中国考试的起源 [J]. 华中师范大学学报（哲学社会科学版），1991，(4)：27－31.

[82] 黎安琪著，石伟平译. 中英教育评定体制的比较研究 [J]. 全球教育展望，1994，(6)：1－8.

[83] 朱利华编译. 美国的 SAT 和 ACT [J]. 比较教育研究, 1995, (2): 55.

[84] 段连运. 国际化学奥林匹克和我国的化学竞赛 [J]. 化学教育, 1995, (2): 39 -44.

[85] 臧铁军. 考试的哲学问题的思考 [J]. 教育研究, 1998, (7): 59.

[86] 汪永钦. 标准化考试及其命题与组卷 [J]. 职业教育研究, 1998, (5): 11 -12.

[87] 康巨珍, 尚玄. 论考试的原则、分类及试卷结构 [J]. 现代技能开发, 1999, (7): 7 -8.

[88] 王毓珣. 国外教育质量监控述评 [J]. 外国教育研究, 1999, (6): 23 -26.

[89] 杨学为. 中国需要"科举学" [J]. 厦门大学学报(哲社版), 1999, (4): 17 -19.

[90] 王天平. 科举考试的历史地位与作用 [J]. 中国高校招生, 1999, (4): 59 -60.

[91] 姜言东. 高考将怎样改革? [J]. 人民教育, 1999, (3): 14 -16.

[92] 张玫欣. 用标准分数进行考试成绩评价 [J]. 教育探索, 2000, (12): 61 -62.

[93] 何明高, 吴福根. 考试分数分布、原始分数与导出分数 [J]. 教育探索, 2000, (3): 21 -22.

[94] 杨维忠, 李彤. 从信度与效度的关系看考试成绩的有效性 [J]. 继续教育研究, 2000, (2): 72 -74.

[95] 钱可铭. 高中会考制度的改革与发展 [J]. 教育评论, 2000, (5): 1 -5.

[96] 臧铁军. 21 世纪高中毕业会考的使命与任务 [J]. 教育研究, 2000, (9): 43 -45.

[97] 郑若玲．高考竞争与科目改革 [J]．高等教育研究，2000，(4)：41 -44.

[98] 陈文，陈华忠．考试与评价制度改革的探索 [J]．教育探索，2001，(10)：47 -48.

[99] 朱信和．以人为本，注重评价的多面性——谈素质教育中的学业评价 [J]．考试 (中考版)，2002，(10)：4.

[100] 刘清华．高考与会考的关系 [J]．教育发展研究，2002，(6)：31 -34.

[101] 杨学为．从高考改革看会考的重要性 [J]．教育研究，2002，(12)：24 -26.

[102] 徐建春．高校招生制度改革的思考 [J]．教育发展研究，2003，(Z1)：135 -137.

[103] 杨光富．当今美英法日四国高考制度 [J]．外国中小学教育，2002，(2)：41 -43.

[104] 董奇，赵德成．发展性教育评价的理论与实践 [J]．中国教育学刊，2003，(8)：41 -45.

[105] 汪茂吾．论中学生考试心理障碍的表现与消除 [J]．教学与管理 (中学版)，2003，(19)：77 -78.

[106] 柯洪霞，李玲．关于小学生考试心理的比较研究 [J]．当代教育科学，2003，(20)：43 -44.

[107] 张莉莉．质性评价的有效尝试：通过学生成长记录袋实现评定的发展性功能 [J]．比较教育研究，2003，(1)：47 -51.

[108] 周军．高考科目设置改革的思考 [J]．湖北招生考试，2004，(104)：24 -29.

[109] 徐静，王瑞烽．标准化测验在理解上的一致与分歧 [J]．中国考试 (研究版)，2004，(1)：24 -26.

[110] 贺斌．关于高中招生制度改革的思考与构思 (上)

[J]. 教育理论与实践, 2004, 24 (5): 21 -25.

[111] 贺斌. 关于高中招生制度改革的思考与构思(下)[J]. 教育理论与实践, 2004, 24 (6): 27 -31.

[112] 高飞. 中国古代考试制度简论 [J]. 教育理论与实践. 2004, 24 (8): 23 -24.

[113] 王兴亚. 孙中山对中国古代考试历史经验的总结 [J]. 湖北招生考试, 2005, (8): 16 -20.

[114] 戴海琦. 考试题库的制作 [J]. 考试研究, 2005, 1 (1): 47 -55.

[115] 陈彬. 对我国高考制度改革的思考 [J]. 当代教育科学, 2005, (24): 50 -52.

[116] 陈燕燕, 彭慧. 论新课程背景下高考制度变革方略 [J]. 现代中小学教育, 2005, (6): 70 -72.

[117] 江琦, 张大均. 中学生考试心理问题及其发展的研究 [J]. 心理科学, 2005, (1): 227 -229.

[118] 江琦, 张大均. 中学生考试心理素质及其发展 [J]. 心理科学, 2005, (4): 841 -844.

[119] 王静韵. 浅谈布鲁姆掌握学习策略对我们的启示 [J]. 化学教育, 2006, (2): 26 -28.

[120] 黄志红. 新课程背景下普通高中综合素质评价研究 [J]. 课程·教材·教法, 2006, (11): 14 -16.

[121] 王小琴. 例谈试卷编制中的几种失误 [J]. 教学与管理, 2006, (2): 77.

[122] 黄海英. 谈考试工作管理的完善 [J]. 文教资料, 2006, (31): 17 -18.

[123] 顾栋明, 周拥军. 学科竞赛团队式管理的尝试 [J]. 文教资料, 2006, (3): 54 -55.

[124] 戴家干. 从考试到评价. 教育改革的时代任务 [J].

中国高等教育，2007，(13－14)：21－23.

[125] 朱行建. 基于标准的科学课程学业考试：一种考试目标的分类框架及应用 [J]. 当代教育科学，2007，(16)：11－14.

[126] 潘懋元. 教育与考试——目的与方法 [J]. 教育与考试，2007，(1)：1.

[127] 申德全. 中考试题中的开放性试题类型及解题策略 [J]. 课程教材教学研究（中教研究），2007，(Z3)：55－56.

[128] 王有鹏. 开放性试题的含义、意义、功能及命题原则 [J]. 教学与管理，2007，(4)：76－78.

[129] 张航. 考试命题计划的研究 [J]. 中国科教创新导刊，2007，(459)：64－65

[130] 杨德平. 杜绝考试作弊的对策分析 [J]. 当代教育科学，2007，(2)：62－63.

[131] 李美华. 试论考试作弊的应对策略与监控措施 [J]. 教育与考试，2007，(5)：22－23.

[132] 王丽. 对建立考试管理标准化系统的思考 [J]. 中国考试，2007，(6)：15－17.

[133] 房源清. 我国普通高中毕业会考制度发展述评 [J]. 考试研究，2007，(2)：22－23.

[134] 唐苏琼. 现行高考制度的弊端及改革对策 [J]. 教育评论，2007，(6)：74－76.

[135] 刘海峰. 高考改革的回顾与展望 [J]. 教育研究，2007，(11)：19－24.

[136] 杨聪聪. 美、日、英中小学学业评价前沿探究 [J]. 世界教育信息，2007，(1)：26－29.

[137] 席珊珊. 综合实践活动课程实施的若干关键问题 [J]. 继续教育研究，2008，(5)：162－163.